



Attachment (1)  
Accession 10/822,303  
Nucleotide

My NCBI [7]  
[Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search  for

[Limits](#) [Preview/Index](#) [History](#) [Clipboard](#) [Details](#)

Format: [GenBank](#) [FASTA](#) [Graphics](#) [More Formats](#) [Download](#) [Save](#) [Links](#)

GenBank AY274119.3

## SARS coronavirus TOR2, complete genome

[Comment](#) [Features](#) [Sequence](#)

LOCUS AY274119 29751 bp RNA linear VRL

24-MAR-2004

DEFINITION SARS coronavirus TOR2, complete genome.

ACCESSION AY274119

VERSION AY274119.3 GI:30248028

KEYWORDS

SOURCE SARS coronavirus Tor2

ORGANISM SARS coronavirus Tor2

Viruses; ssRNA positive-strand viruses, no DNA stage;

Nidovirales;

Coronaviridae; Coronavirinae; Betacoronavirus.

REFERENCE 1 (bases 1 to 29751)

AUTHORS He, R., Dobie, F., Ballantine, M., Leeson, A., Li, Y.,

Bastien, N.,

Cutts, T., Andonov, A., Cao, J., Booth, T. F., Plummer, F. A.,

Tyler, S.,

Baker, L. and Li, X.

CONSRM BCCA Genome Sciences Centre, British Columbia Centre

for Disease

Control and National Microbiology Laboratory Canada

Analysis of multimerization of the SARS coronavirus

nucleocapsid

protein

JOURNAL Biochem. Biophys. Res. Commun. 316 (2), 476-483 (2004)

PUBMED 15200242

REFERENCE 2 (bases 1 to 29751)

CONSRM BCCA Genome Sciences Centre, British Columbia Centre

for Disease

Control and National Microbiology Laboratory Canada

Direct Submission

JOURNAL Submitted (13-APR-2003) Genome Sciences Centre, British

Columbia

Vancouver, BC V5Z

4E6, Canada

REFERENCE 3 (bases 1 to 29751)

CONSRM BCCA Genome Sciences Centre, British Columbia Centre

for Disease

Control and National Microbiology Laboratory Canada

Direct Submission

JOURNAL Submitted (23-APR-2003) Genome Sciences Centre, British

Columbia

Vancouver, BC V5Z

4E6, Canada

REMARK Sequence update by submitter

REFERENCE 4 (bases 1 to 29751)

CONSRM BCCA Genome Sciences Centre, British Columbia Centre

for Disease

Control and National Microbiology Laboratory Canada

[Change Region Shown](#)

[Customize View](#)

[Analyze This Sequence](#)

[Run BLAST](#)

[Pick Primers](#)

### SARS Coronavirus Resource

Compiled data and information relevant to SARS coronavirus.

[Recent activity](#)

[Turn Off](#) [Clear](#)

Laboratory safety aspects of SARS at Biosafety Level 2.

SARS coronavirus TOR2, complete genome

SARS coronavirus TOR2, complete genome

SARS coronavirus TOR2, complete genome

AY274119 (1) [Nucleotide](#)

[» See more.](#)

### All links from this record

[Related sequences](#)

[Full text in PMC](#)

[Genome](#)

[Other nucleotides for species](#)

[Probe](#)

[Protein](#)

[PubMed](#)

TITLE Direct Submission  
 JOURNAL Submitted (30-APR-2003) Genome Sciences Centre, British Columbia  
 Vancouver, BC V5Z 4E6, Canada  
 REMARK Sequence update by submitter  
 COMMENT On Apr 30, 2003 this sequence version replaced  
 gi:30088476.

PubMed (weighted)  
 RefSeq nucleotide for species  
 Taxonomy

FEATURES  
 source  
 Location/Qualifiers  
 1..29751  
 /organism="SARS coronavirus Tor2"  
 /mol\_type="genomic RNA"  
 /isolate="Tor2"  
 /isolation\_source="patient #2 with severe acute respiratory syndrome (SARS)"  
 /db\_xref="taxon:247964"  
 /country="Canada: Toronto"  
 join(265..13392,13392..21485)  
 /ribosomal\_slippage  
 /codon\_start=1  
 /product="replicase 1AB"  
 /protein\_id="AAP41036.1"  
 /db\_xref="GI:30795144"  
 /translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRGFGD"

SVEEALSEARE

HLKNGTCLGLELEKGVLPQLEQFPYVPIKRSDALSTNHGHKVVELVASMDDGIQYGRSGI  
 TLGLVLPVHVGTFPIAYENVLLRKNGKAGGHSYGIDLKSYDLGDELGTDPIEDYEQN  
 WNTKHGSGALRELTRELNGGAVTRYVDNFCGPDGYPDLCIKFLARAGKSMCTLSEQ  
 LDYIESKRGVYCCRDHEHEIAWFTERSDKSYEHQTPFEIKSAKKFDTPFGCECPKFVPF  
 LNSKVKVIQPRVEKKTEGFMGRISVYPVASPQRCNNMHLSTLMKCNHCEVSWQTC  
 DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVMPCPACQDPEIGPEHSVADYHNHNS  
 IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPASADIGSGHTGITGDNVETLNEDL  
 LEILSRERVNINIVGDFHLEEVAILASFSASTSAFIDTIKSLDYKSPKTI VESCGN  
 YKVTGKGVKAWNIGQQRSLVTPLCGFPSPQAAGVIRISIFARTLDAANHISIPDLQRAA  
 VTILDGISEQSLRLVDAMVYTSDLLTNSVIMAYVTGGLVQQTQWLSNLLGTTVEKL  
 RPIFEWIEAKLSAGVEFLKDAWEILKPLITGVDFIVKGQIQVASDNIKDCVKCFIDVV  
 NKALEMCIQVTIAGAKRLSLNIGEVPIAQSCKGLYRQCIRGKEBQLQLMPLKAPKEVT  
 FLEGDSDHTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTIVCVNGLMLLEIKDKQE  
 YCALSPGLLATNNVFLKGGAPIKGVTFGEDITVNEVQGYKNVRIITFELDERVDKVLNE  
 KCSVYTVESGTEVTEFACVVAEAVVKTLQFVSDLLTNMGIDLDENSVATFYLFDDAGE  
 ENFSRMYCSFPYDDEEEDDABCEEBIDEETCEHEYGTEDDYQQLPLEFGASAEIVR  
 VEEEEEEDWLDOTTEQSEIEPEPEPTPEEPVNPQTYGLKLTNDVAIKCVDIVKEAQA  
 NPMVIVNAANIHLKHGGGVAGALNKATGAMQKESDDYIKLNGPLTVGSGCLLSGHN  
 AKKCLHVVGPNLNAGEDIQLLKAAENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT  
 VRTQVYIAVNDKALYEQVMDYLDNLKPRVEAFKQEPPEPTEDSKTEKSVVQKFPVD  
 KPKIKACIDEVTTTLEETKPLTNKLLLPADINGKLYHDSQNMRLRGEDMSFLEKDAPYM  
 VGDVITSGDITCVVIPSKKAGGITTEMLSRALKKVPVDEYITTPYGGQAGYTTLEAKT  
 ALKKCKSAFYVLPSEAPNAKEEILGTVSNLREMLAHAEETRLKMPICMDVRAIMATI

QRKYKGIKIQEGIVDVGVRFFYTSKEPVASIITKLNSLNEPLVMTPIGVYTHGFNLE  
EAARCMRSLKAPAVVSVSSPDAVVTYNGYLTSSSKTSEEHFVETVSLAGSYRDWSYSG  
QRTELGEVFLKRGDKIVYHTLESPEVPHLDGEVLSLDKLSLSLREVKTIKVFTITVD  
NTNLHTQLVDMSTMTYGGQFGPTYLDGADVTKIKPHVNHEGKTFVFLPSDDTLRSEAFE  
YYHTLDESFGLGRYMSALNHTKKWKFPQVGGTSLIKWADNNCYLSVLLJALQLEVKFN  
APALQEAYYRARGDAANPCALILAYSNTKTVGELGDVRETMTHLQHANLES AKRVLN  
VVCKHCGQKTTTLTGVEAVMYMTLSYDNLKTGVSIPVCGRDATQYLVQCESSFVMM  
SAPPAEYKLQGGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD  
VFYKETS YTTTIKPVSYKLDGVYTEIEPKLDGYKKDNAYYTEQPIDLVPITPLPNA  
SFDNFKLTCNTRKFAADDLNQMTGFTKPSARELSVTFPDNLNGDVVAIDYRHYSASFCK  
GAKLLHKPIVMHINQATTKTTFKPNTWCLRLWSTKPVDTSSNSFEVLAVEDTQGMNLI  
ACESQQPTSSEVVENPTIQKEVIECDVKTTEVGVNVLKPSDEGVKVTQELGHEDLMA  
AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPMWKILAYVKPFLQAAAITTSN  
CAKRLAQRVFNMYPVVFTLLFQLCTFTKSTNSRIASLPTTIAKNSVKSAKLCLDA  
GINYVKS PKFSKLF TIAMWLLLSICGLSLICVTAAGVLLSNFGAPSYCNGVRELYL  
NSSNVTTMDFCGSGFPČSICLSGLDSDSYPALETIQVTISSYKLDLTILGLAAEWVL  
AYMLFTKFFYLLGLSAIMQVFFGYFASHFISNSWLMWFIISIVQMAPVSAMVRMYIFF  
ASFYIWKSYVHTMDGCTSTCMCYKRNATREVECTTIVNGMKRSFYVYANGGRGFC  
KTHWNCLNCDTFCGTGTFISDEVARDSLQFKRPINPTDQSSYIVDSVAVKNGALHL  
YFDKAGQKTYERHPLSHFVNLDNLRANNTKGLSPINIVVFDGSKCODESASKSASVY  
SQIMCQFILLLDQALVSDVGDSTEVSVMFDAYVDTFSATFSVPMEKLKALVATAHSE  
LAKGVALDGVLSIFVSAARQGVVDVDTKDVI ECLKLSHSDLEVTGDCNNFMITY  
NKVENMTPRDLGACIDCNARHINAQVAKSHVSLIWNKDYMSLSBQLRKQIRSAAKK  
NNIPFRLTCATTRQVNVITTKISLKGKIVSTCFKMLKATLLCVLAAALVCYIVMPV  
HTLSIHDGYTNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDAMFSGRQGSYKNDKSC  
FVVAIIITREIGFIVPGLPGTVLRAINQDGLHFLPRVFSAVGNICYTPSKLIEYSFDA  
TSACVLAAECTIFKIDAMGKVPVYCYDNLLEGSISYSELRPDTRYVLMDSIIQFPNT  
YLEGSVRVVTTFDAEYCRHGTCESEVGICLSTSGRWLVNNEHYRALSGVPCGV DAMN  
LIANIPTLVQPVALDVSAVVAGGIIAILVTCAAYYFMKFRVFGBYNHRVVAANAL  
LFLMSFTILCLVPAYSPLPGVYSVFYLYLTYFTNDVSLAHLQWFMFSPVPPFWIT  
AIYVFCISLKHCHWFFNNYLKRVMPNGVTFSTPEAAALCTFLNKEMYLKLRETLIL  
PLTQYNRYLALYNKYKYPGALDTSYREACCHLAKALNDFSNGADVLYQPPQTSI  
TSAVLQSGFRKMAFSPGKVEGCMVQVTCGTTTLNGLWLDOTVYCPRHIVICTAEDMLNP  
NYEDLLIRKSNHSFLVQAGNVQLRVIGHMSQNCLLRLKVDTSNPKTPKYKPVRIQPGQ  
TFSVLACYNGSPSGVYQCAMRPNHTIKGSPNGSCGSGVGNIDYDCVSPCYMHIMELP  
TGVHAGTDLEGKPYGPFVDRQTAQAAGTDTTITLNLWLWLYAAVINGDRWPLNRPTTT

Write to the Help Desk  
NCBI | NLM | NIH  
Department of Health & Human Services  
Privacy Statement | Freedom of Information  
Act | Disclaimer

LNDFNLVAMKYNIEPLTQDHDVILGPLSAQTGIADVLMCAALKELLQNGMNGRTILGS  
TILEDEFTPFDDVVRQCSGVTFQGGPKIKVGTTHWMLLTFTLSLLVQSTQNSLFFP  
VYENAFLPFTLGIMAIACAMLLVKHEAFLCLFLLPSLATVAYFNMVMPASWVMRI  
MTWLEADTSLSGYRLKDCVNYASALVLLILMTARTVYDDAARRVWIMNVITLVYKV  
YYGNALDQAISMALVISVTSNYSYGVVTTIMFLARAIVFCVEEYPLLFITGNTLQCI  
MLVYCFPLGYCCCCYFGLFCLLNRYPRITLGVYDVLVSTQEFRYMNSQGLLPKSSIDA  
FKLNIKLLGIGGKPKCIKVATVQKMSDVKCTSVLLSVLQQLRVESSEKILWAQCVQLH  
NDILLAKDTTEAFKMSVLLSVLLSMQGAVDINRLCEMLDNKRATLQATASEFSSLPS  
YAAAYATAQEAYEQAVANGDSEVLLKKLKSINLVAKSEFDRDAAMQRKLEKMDQAMTQ  
MYKQARSEDKRAKVTSAMQTMFLTMLRKLDNDALNNIINNARDGCVPLNIIPLTAAK  
LMVVVVDYGTYYKNTCDGNTFTYASALWEIQQVVDADSKIVQLSEINMDSNPLAWPLI  
VTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTDDNALAYNNNSKGRFVLALL  
SDHQDLKWARFPKSDGTGTIYTELEPPCRFVIDTPKGPVKYLYFIKGLNNLNRGMVL  
GSLAATVRLQAGNATEVPANSTVLSFCAVDPKAYKYDLASGGQPIITNCVKMLCTH  
TGTGQAITVTPCANMQDESFGGASCCLYCRCHIDHPNPKGFCDLKGKYVIPITTCAND  
FVGFTLRNTVCTVCGMWKGYGSCDQLREPLMQSADASTFFKRVCGVSAARLTPCGTG  
TSTDVVYRAFDIYNEKVAGFAFLKTNCCRQKQDEGNLLDSYFVVKRHTMSNYQHE  
ETIYNLVKDCPAVAVHDFKFRVDGDMVPHISRQLTKYTMADLVYALRHFDEGNCDT  
LKEILVTYNCCDDYFNKDWYDFVENPDILRVYANLGERVRSLLKTVQFCDAMRDA  
GIVGVLTLDNQDLNGMWDFDGVQVAPGCGVPIVDSYYSLLMPILTLTRALAABESH  
DADLAKPLIKWDLKYDFTBERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCANFV  
LFSTVPFPPTSFGPLVRKIFVDGVFPVVSIGYHRELGVVHNQDVLNHSRSLSKELLV  
YAADPAMHAASGNLLDKRTTCFSAALTNVAPQTVKPGNFNDPYDFAVSKGPFKE  
GSSVELKHFFFAQDGNAAISDYDYRYNLPMTCDIRQLLFVVEVDKYFDYDGGCIN  
ANQVIVNMLDKSAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITITQNLKYA  
ISAKNRARTVAGVSICSTMTRQPHQKLLKSIATRGAATVIGTSKFYGGWHMLKTV  
YSDEVTPHLMGWDPKCDRAMPNMLRIMASLVLRKHNTCCNLSHRFYRLANECQVL  
SEMVMCGGSLYVKPGTSSGDATTAYANSVFNICQAVTANVALLSTDGNKIADKYVR  
NLQHRLYECLYRNRDVDHEFVDEFYAYLRKHPSSMTLSDDAVVCYNSNYAAQGLVASI  
KNFKAVLYYQNNVPMSEAKNCTETDLTKGPHFCQHTMLVKQDDYVYLPYDPSRI  
LGAGCFVDDIVKTGTLMIERFVSLAIDAYPLTKHPNQEYADVPHLYLQYIRKLHDEL  
TGHMLDMYSVMLTNDNTRSYNEPEFYEAMYPHTVLQVAGACVLCNSQTSRLCGACIR  
RPFLLCKCCYDHVISTSHKLVLVSNPYVCNAPGCDVTDVTLQYLGMSYCYCKSHKPP  
SFPLCANGQVFLYKNTCVGSDNVTFNAIATCDWTNAGDYILANTCTERLKLFAET  
LKATEETPKLSYGIATVREVLSDRELHLSWEVGKPRPPLNRNYVFTQYRVTKNSKVQI  
GEYTFEKGDYCDVYVRGTTTYKLVNGDYFVLTSTHTVPLSAPTLVPQEHYVRITGLY

```

PTLNISDEFSSNVANYQKVMQKYSTLQGPPGCTGKSHFAGLALYYPSARIVYTACSH
AAVDALCEKALKYLPIDKCSRIIPARARVECFDFKFNSTLBQYVFCVNALPETTAD
IVVFDEISMATNYDLSSVNARLRAKHVYVYIGDPAQLPAPRTLLTKGLTEPEYFNSVCR
LMKTIIGPDMFLGTCCRRCPAEIVDTVSALVYDNKLKAHKDKSAQCFKMFYKGVITHDVS
SAINRPQIGVVREFLTRNPAMRKAVFISPYNSQNAVASKIILGLPTQTVDSQSGSEYDY
VIFTQTETTAHSCNVRNFVAITRAKIGILCIMSDDRLYDKLQFTSLEIPRRNAVTLQ
AENVTLGLFKDCSKIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGIPKDMTYRRLISMM
GFMNYQVNGYPNMFITREEAIRHVRWVIGFDVEGCHARTDAVGTNLPLQLGFSTGVN
LVAVPTGQVDTENNTEFTRVNAKPPPGDQFKHLIPLMYKGLPWNVRIKIVQMSDTL
KGLSDRVVFLVAHGFEITSMKYFVIGPERTCCLCDKRACTFSTSSDTYACWNHSGV
FDYVYNPFMIDVQWQFTGNLQSNHDDHCQVHGNAHVASCDAIMTRCLAVHECFVKRV
DWSVEYPIIGDELVRNSACRQVQHMVVKALLADKFFVLHDIGNPKAICKVPQAEVEW
KFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCVNDRYPANAIVCRPDRVL
SNLNLPGCGDGLSYVNHAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVVSDD
YVPLKSATCITRNLGGAVCRHHANEYRQYLDAYNMMSAGPSLWIYKQFDTYNLWNT
FTRLQSLNVAYNVVKGHFDGHAGEAPVSIINNAVYTKVDGIDVEIFENKTTLPVNV
AFELWAKRNKIPVPEIKIINNLGVDIAANTVINDYKREAPAHVSTIGVCTMTDIAKKP
TESACSSLTVLFDGRVEGQVDFRNARNGVLLITEGSKVLTPSKGPAQASVNGVTLIG
ESVKTQFNYPKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETDFLELANDEFIQRKYL
BGYAFEHIVYDPSHGQLGLHLMIGLAKRSQDPLKLEDFIPMDSTVKNYFITDAQT
GSSKCVCSVIDLLDDFVEIISKQDLSVISKVKVTIDYAEISFMLWCKDGHVETFPY
KLQASQAMQPGVAMFNLYKMQRMLLEKCDLQNYGENAVIPKGIMMVAKYTLQCLQYLN
TLTLAVPYNMRVHFGAGSDKGAVPTAVRLQWLPTGTLVDSDLNDFVSDADSTLIG
DCATVHTANKWDLIISDMYDPRTKHVTKENDSKEGFFTYLCGFIKQKLALOGSIAVKI
TEHSWNADLYKLMGHFSWTAFTVNVNASSSEAFLLIGANYLKGPKQIDGYTMHANYI
FWRNTNPQLSSYSLFDMSKFPLKRGTAVMSLKENQINDMIYSLLEKGRIIIRENRR
VVSSTILVNN"
21492..25259
/codon_start=1
/product="spike glycoprotein"
/protein_id="AA033714"
/db_xref="GI:30795145"
/translation="*MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQH
TSSMRGVVYPD
EIFRSDTLYLTQDLFLPFYSNVITGPHITNHTPGNPVIFPKDGIYFAATEKSNVVRGNV
FGSTMNNKSQSVIIINNSTNVVIRACNFELCDNPPFAVSKPMCTQHTMIFDNARFNT
PEYISDAFSLDVSEKSGNFKHLREFVFNKNDGFLYVYKGYQPIDVVRDLPSGFNTLKP
IFKLPLGINITNFRAILTAFSPAQDWTGTSAAAYFVGYLKPITPMLEYDENGTIIDAV
DCSQNPALAEKCSVKSFEDKGIYQTSNFRVVPDGVVRFPNITNLCPFGEVFNATKF
PSVYANERKKIISNCVADYSVLYNSTFFSTFKCYGVSAKLNLDLCFSNVYADSPVVKGD

```

<http://www.ncbi.nlm.nih.gov/puccore/30248028>

LYI IKLVFLWLLWVTLACFLAAVYRINNVITGGIAIACIVGLMWLSYFVASRFLF  
 ARTSMWSFNPETNILLNVPLRGITVTRPLMESELVIGAVIIRGLHRMAGHSLSGRCDI  
 KDLPKETIVATSRSLSYKLGASQVRVGTDSGFAAYNRYRIGNYKLNTHAGSNDNIAL  
 LVQ"  
 27074..27265  
 /codon\_start=1  
 /product="Orf7"  
 /protein\_id="AAP41042.1"  
 /db\_xref="GI:30795150"  
 /translation="MFHLVDPQVTIAEILIIIMRTFRIAIWNLDVII  
 SSIVRQLFKPL  
 TKNYSELDDDEPEMELDYP"  
 27273..27641  
 /codon\_start=1  
 /product="Orf8"  
 /protein\_id="AAP41043.1"  
 /db\_xref="GI:30795151"  
 /translation="MKIILFLTLIVFTSCELYHYQBCVRGTTVLLEK  
 PCPSGTYEGNS  
 PFHPLADNKFALCTCTSTHFAACADGTRHTYQLRARSVSPKLFIRQSEVQQLYSLF  
 LIVAALVFLILCFTIKRTE"  
 27638..27772  
 /codon\_start=1  
 /product="Orf9"  
 /protein\_id="AAP41044.1"  
 /db\_xref="GI:30795152"  
 /translation="MNELTLIDFYLCFLAFLFLVLIMLIIFWFSLE  
 IQDLEEPCTKV"  
 27779..27898  
 /codon\_start=1  
 /product="Orf10"  
 /protein\_id="AAP41045.1"  
 /db\_xref="GI:30795153"  
 /translation="MKLLIVLTICISLCSICTTVQRCASNKPHVLED  
 PCKVQH"  
 27864..28118  
 /codon\_start=1  
 /product="Orf11"  
 /protein\_id="AAP41046.1"  
 /db\_xref="GI:30795154"  
 /translation="MCLKILVRYNTRGNTYSTANLWALGKVLPPHRW  
 HTMVQCTCPNV  
 TINQDPAGGALIAACWYLHEGHQTAARFDVLVLNKRNT"  
 28120..29388  
 /codon\_start=1  
 /product="nucleocapsid protein"  
 /protein\_id="AAP41047.1"  
 /db\_xref="GI:30795155"  
 /translation="MSDNGPQSNQRSAPRITFGGPTDSTDNQNGGR  
 NGARPKQRRPQ  
 GLPNNTASWFTALTQHCKEELRPPRGQGVPIINTSGPDDQIGYRRATRRVRGGDGKM  
 KELSFRWYPYLLGTGPEASLPYGANKGIVWVATEGALNTPKDHIGTRNFRNNAATVL  
 QLPQGTTLPGKGYAEGSRGSGQASSRSSRSRGRNSRSTPGSSRGNSPARMASGGGET  
 ALALLLLDLRLNQLSKVSGKGQQQQQQTVTTKSAEASKKPRQKRTATKQYNVTQAPG  
 RRGPEQTQGNFGDQLIRQGTDYKHWPQIAQFAPSASAFPGMSRIGMEVTPSGTWLTY  
 HGAIKLDDKDPQFKDNVILLNKHIDAYKTTPPTEPKDKKKKKTDEAQLPQPKKQPT  
 VTLLPADMDDFSRLQLNSMSGASADSTQA"  
 28130..28426  
 /codon\_start=1  
 /product="Orf13"  
 /protein\_id="AAP41048.1"  
 /db\_xref="GI:30795156"  
 /translation="MDPNQTNVVPALHLVDPQIQLTITRMEDAMGQ

GQNSADPKVYP

IIILRLGSQLSLMARRNLDLEARAFQSTPIVVQMTKLATTEELPDEFVVTAK\*  
 28583..28795  
 /codon\_start=1  
 /product="Orf14"  
 /protein\_id="AAP041040.1"  
 /db\_xref="GI:30795157"  
 /translation="MLPPCYNFLKEHQCKASTQREAAAVKPLIAP"

HHVVAVIQEIQ

LLAAVGEIILLEWLAIEVVKLPSRYCC"

ORIGIN

```

1 atattagggtt tttacctacc caggaaaagc caaccaacct cgaatctcttg
tagatctgtt
61 ctctaaacga accttaaaat ctgtgtagct gtcgctcggc tgcagtccta
gtgcacctac
121 gcagataaaa caataataaa ttttactgtc gttgacaaga aacgagtaac
tcgtccctct
181 tctgcagact gcttacgggt tcgtccgtgt tgcagtcgat catcagcata
cctagggttc
241 gtccgggtgt gaccgaaagg taagatggag agcctgttc ttggtgtcaa
cgagaaaaaa
301 cagctccaac tcagtttgcc tgccttcag gttagagacg tgcagtgtcg
tggcttcggg
361 gactctgtgg aagaggccct atcggaggca cgtgaacacc tcaaaaaatg
cacttctggt
421 ctagttagac tggaaaaagg cgtactgcc cagcttgaa agccctatgt
gttcattaaa
481 cgttctgatg ccttaagcac caatcacggc cacaaggctg ttgagctggt
tgcagaaaatg
541 gacggcatc agtacggctg tagcgggata acactgggag tactcgtgcc
acatgtgggc
601 gaaaccccaa ttgcataccg caatgttctt ctctglaaga acggtaataa
gggagccggt
661 ggctcatagc atggcatcga tctaaagctc tatgacttag gtgacgagct
tggcactgat
721 ccatttgaag attatgaaca aaactggaac actaagcatg gcagtgtgtc
actcgtgaa
781 ctactcgtg agctcaatgg aggtgcagtc actcgtctat tgcacaacaa
ttctgtggc
841 ccagatgggt accctcttga ttgcatcaa gatttctcgc cagcgcggg
caagtcaatg
901 tgcactcttt ccgaacaact tgattacatc gagtcaaga gaggtgtcta
ctgctgccgt
961 gaccatgagc atgaaattgc ctgggttact gagcgtctctg ataagagcta
cgagcaccag
1021 acacccctcg aaattaagag tgccaagaaa ttgacactt tcaaggggga
atgcccacaa
1081 tttgtgttcc ctcttaactc aaaagtcaaa gtcattcaac cacygtgtga
aaagaaaaag
1141 actgagggtt tcatggggcg tatacgtctc gtgtaccctg ttgcatctcc
acaggaggtt
1201 aacaatatgc acttgtctac ctgtatgaaa tgaatcatt gcgagtgaag
ttcatggcag
1261 acgtgcgact ttctgaaagc cacttgtgaa cattgtggca ctgaaaattt
agttattgaa
1321 ggacctacta catgtgggta cctacctact aatgctgtag tgaanaatgcc
atgtcctgcc
1381 tgtcaagacc cagagattgg acctgagcat agtgttcgag attatcaca
ccactcaaac
1441 attgaaactc gactccgcaa gggaggtagg actagatgct ttggaggctg
tgcgtttgcc
1501 tatgttggtc gctataataa gcgtgcctac tgggttcctc gtgctagtgc
tgatattggc
1561 tcaggccata ctggcattac tggtagacaa gtggagacct tgaatgagga
tctccttgag
1621 atactgagtc gtgaacgtgt taacattaac attgttggcg attttcattt
gaatgaagag
1681 gttgccatca ttttggcacc ttctctgtct tctacaagtg cttttattga
cactataaag
1741 agtcttgatt acaagtcttt caaaaccatt gttgagtcct gcggttaacta
taaagtacc
```



1801 aagggaaagc ccgtaaaagg tgcctgggac attggacaac agagatcagt  
 tttaacacca  
 1861 ctgtgtggtt ttccctcaca ggctgctggt gttatcagat caatttttgc  
 ggcacacatt  
 1921 gatgcagcaa accactcaat tcttgatttg caaagagcag ctgtcaccat  
 acttgatggt  
 1981 atttctgaac agtcattacg tcttgctgac gccatgggtt ataactcaga  
 cctgctcacc  
 2041 aacagtgtca ttattatggc atatgtaact ggtgggtctg tacacacagac  
 ttctcagtgg  
 2101 ttgtctaate ttttgggcac tactgttgaa aaactcaggc ctatctttga  
 atggattgag  
 2161 gcgaactta gtgcaggagt tgaattcttc aaggatgctt gggagattct  
 caaatttctc  
 2221 attacaggtg tttttgacat cgtcaagggt caaatacagg tctgttcaga  
 taacatcaag  
 2281 gattgtgtaa aatgcttcac tgatgttgtt aacaaggcac tcgaagtgtg  
 cattgatcaa  
 2341 gtcactatcg ctggcgcaaa gttgcgatca ctcaacttag gtgaagctct  
 catgctcaa  
 2401 agcaaggagc ttaccgtcac gtgtatacgt ggcaaggagc agctgcaact  
 actcatgctt  
 2461 cttaaggcac caaagaagt aacctttctt gaaggtgatt cacatgacac  
 agtacttacc  
 2521 tctgaggagg ttgttctcaa gaacggtgaa ctggaagcac tcgagacgcc  
 cgttgatagc  
 2581 ttcacaaatg gagctatcgt tggcacacca gtctgtgtaa atggcctcat  
 gctcttagag  
 2641 attaaggaca aagaacaata ctgcgcattg tctcctgggt tactggctac  
 aaacaatgct  
 2701 tttcgtctaa aaggggggtg accaattaaa ggtgtaacct ttggagaaga  
 tactgtttgg  
 2761 gaagttcaag gttacaagaa tgtgagaatc acatttgagc ttgatgaacg  
 tgttgacaaa  
 2821 gtgcttaagt aaaagtgtct tgtctacact gttgaatccg gtaccgaagt  
 tactgagttt  
 2881 gcatgtgttg tagcagaggc tgtgtggaag actttacaac cagtttctga  
 tctccttacc  
 2941 aacatgggta ttgatcttga tgagtggagt gtactacat tctacttact  
 tgatgatgct  
 3001 ggtgaagaaa acttttctac acgtatgtat tgttctcttt accctccaga  
 tgagggaaga  
 3061 gaggaagcag cagagtgtga ggaagaagaa attgatgaaa cctgtgaaca  
 tgagtacggg  
 3121 acagaggatg attatcaagg tctcctctg gaatttgggt cctcagctga  
 aacagttcga  
 3181 gttgaggaag aagaagagga agactggctg gatgatacta ctgagcaatc  
 agagattgag  
 3241 ccagaaccag aacctacacc tgaagaacca gttaatcagt ctactgggta  
 tttaaaactt  
 3301 actgacaatg ttgccattaa atgtgtgac atcgtaaggg aggcacaaag  
 tgctaactct  
 3361 atgggtgattg taaatgctgc taacatacac ctgaaacatg gtgtgggtgt  
 agcaggtgca  
 3421 ctcaacaagg caaccaatgg tgccatgcaa aaggagagtg atgattacat  
 taagctaaat  
 3481 ggccctctta cagtaggagg gtctgttttg ctttctggac ataactctgc  
 taagaagtgt  
 3541 ctgcatgttg ttgacactaa cctaataatg ggtgaggaca tccagcttct  
 taaggcagca  
 3601 tatgaaaatt tcaattcaca ggacatctta ctgacccat tgtgtgacg  
 aggcataatt  
 3661 ggtgctaaac cacttcagtc ttacaagtg tgcgtgcaga cgggtctgac  
 acaggtttat  
 3721 attgcagtca atgacaaagc tctttatgag caggtgtgca tggattatct  
 tgataaacctg  
 3781 aagcctagag tgaagcacc taaacaagag gagccacca acacagaaga  
 tcccaaaact  
 3841 gaggaagaat ctgtcgtaca gaagcctgtc gatgtgaagc caaaaattaa  
 ggctctgcat  
 3901 gatgaggtta ccacaacact ggaagaaact aagtttctta ccaataagtt  
 actctgtttt

3961 gctgatatca atggttaagct ttaccatgat tctcagaaca tgcttagagg  
tgaagatatg  
4021 tcttctcttg agaaggatgc accctacatg gtagggtgat ttatcactag  
tgggtatatac  
4081 acttggtgtg taataccctc caaaaaggct ggtggcacta ctgagatgct  
ctcaagagct  
4141 ttgaagaaag tgccagtta tgagtatata acccgtacc ctggacaagg  
atgtgctggt  
4201 tatacacttg aggaagctaa gactgctctt aagaaatgca aatctgcatt  
ttatgtacta  
4261 ccttcagaag cacctaagtc taaggagag atcttaggaa ctgtatctgc  
gaatttgaga  
4321 gaaatgcttg ctcatgctga agagacaaga aaattaatgc ctatatgcat  
ggatgttaga  
4381 gccataatgg caaccatca acgtaagtat aaaggaatta aaattcaaga  
gggcatcggt  
4441 gactatgggt tccgattctt cttttatact agtaagagac ctgtagcttc  
tattattaag  
4501 aagctgaact ctctaaatga gccgcttgtc acaatgccaa ttggttatgt  
gacacatggt  
4561 tttaatcttg aagaggctgc gcgctgtatg cgttctctta aagctcctgc  
cgtagtgtca  
4621 gtatcatcac cagatgctgt tactacatat aatggatacc tcaactctgc  
atcaaaagaca  
4681 tctgaggagc actttgtaga aacagtttct ttggctggct cttacagaga  
ttggtcctat  
4741 tcaggacagc gtacagagtt aggtgttgaa tttcttaagc gttggtgaca  
aatgtgttac  
4801 cacactctgg agagccctgc cgagtttcat cttgacgggt aggttcttct  
actgacaaa  
4861 ctaaaagagtc tcttatccct gcgggaggtt aagactataa aagtgttcac  
aactgtggac  
4921 aacactaatc tccacacaca gcttggtgat atgtctatga catatggaca  
gcatgttggt  
4981 ccaacatact tggatgggtc tgatgttaca aaaattaac ctcatgtaaa  
tcatgagggt  
5041 aagactttct ttgtactacc tagtgatgac acactacgta gtgaagcttt  
cgagtactac  
5101 catactcttg atgagagttt tcttggtagg tacatgtctg ctttaaacca  
cacaagaaa  
5161 tggaaatttc ctcaagttgg tggtttaact tcaattaat gggctgataa  
caattgttat  
5221 ttgtctagtg ttttattagc acttcaacag cttgaagtca aattcaatgc  
accagcactt  
5281 caagaggctt attatagagc ccgtgctggt gatgctgcta acttttctgc  
actcactac  
5341 gcttacagta ataaaactgt tggcgagctt ggtgatgtca gagaaactat  
gaccatctt  
5401 ctacagcatg ctaatttgga atctgcaag cgagttctta atgtgggtg  
taaactatgt  
5461 ggtcagaaaa ctactacctt aacgggtgta gaagctgtga tgtatatggg  
tactctatct  
5521 tatgataatc ttaagacagg tgtttccatt ccatgtgtgt ggtgctgtga  
tgctacacaa  
5581 tatctagtac aacaagagtc ttctttgtt atgatgtctg caccacctgc  
tgagtataaa  
5641 ttacagcaag gtacattctt atgtgcgaat gagtacactg gtaactatca  
gtgtggtcat  
5701 taactcataa taactgctaa ggagaccctc tatcgtattg acggagctca  
ccttcaaaag  
5761 atgtcagagt acaaaggacc agtgactgat gttttctaca aggaaacatc  
ttacactaca  
5821 accatcaagc ctgtgtcgtg taaactcgat ggagttactt acacagagat  
tgaaccaaaa  
5881 ttggatgggt attataaaaa ggataatgct tactatacag agcagcctat  
agaccttgta  
5941 ccaactcaac cattacaaaa tgcgagtttt gataatttca aactcacatg  
ttctaacaca  
6001 aaatttgctg atgatttaaa tcaaatgaca ggcttcacaa agccagcttc  
acgagagcta  
6061 tctgtcacat tcttccaga cttgaatggc gatgtagtgg ctattgacta  
tagacactat

6121 tcagcgagtt tcaagaaagg tgctaaatta ctgcataagc caatgtgtt  
 gcacattaac  
 6181 cagggtacaa ccaagacaac gttcaaacca aacacttggt gtttactgtt  
 tctttggagt  
 6241 acaaaagccag tagatacttc aaattcattt gaagttcttg cagtagaaga  
 cacacaagga  
 6301 atggacaatc ttgcttggtg aagtcaacaa cccacctctg aagaagtagt  
 ggaataatcct  
 6361 accatacaga aggaagtcac agagtgtgac gtgaaaacta ccgaagtgtg  
 aggaatgttc  
 6421 atacttaaac catcagatga aggtgtttaa gtaacacaag agttaggtca  
 tgaggatctt  
 6481 atggctgctt atgtggaaaa cacaagcatt accattaaga aacctaatag  
 gctttcacta  
 6541 gccttaggtt taaaaacaat tgccactcat ggtattgctg caattaatag  
 tgttccttgg  
 6601 agtaaaattt tggcttatgt caaacatttc ttaggacaag cagcaattac  
 aacatacaat  
 6661 tgcgctaaga gattagcaca acgtgtgttt aacaattata tgccttatgt  
 gtttacatta  
 6721 ttgtccaat tgtgtacttt tactaaaagt accaattcta gaattagagc  
 ttcactacct  
 6781 acaactattg ctaaaaatag tgtaagagt gttgctaata tatgtttgga  
 tgcggcatt  
 6841 aattatgtga agtcacccaa attttctaaa ttgttcacaa tcgctatgtg  
 gctatgtgtg  
 6901 ttaagtattt gcttaggttc tctaactctg gtaactgctg cttttgggtg  
 actcttatct  
 6961 aattttgggt ctccttctta ttgtaatggc gtagagaat tgtatcttaa  
 ttgcttaac  
 7021 gttactacta tggatttctg tgaaggttct tttccttgca gcatttgggt  
 aagtggtat  
 7081 gactcccttg atcttatccc agctcttgaa accattcagg tgacgatttc  
 atcgtacaag  
 7141 ctgacttga caattttagg tctggcgcct gagtgggttt tggcatatat  
 gttgttcaca  
 7201 aaattcttct atttattagg tctttcagct ataactcagg ttttctttgg  
 ctatttctgt  
 7261 agtcatttca tcagcaattc ttggctcatg tggtttatca ttagatttgg  
 acaaatggca  
 7321 cccgtttctg caatgggttag gatgtacac tcttttgcct ctttctacta  
 catatggaag  
 7381 agctatgttc atatcatgga tgggtgcacc tcttcgactt gcattgatgtg  
 ctataagcgc  
 7441 aatcgtgcca cagcgcgtga gtgtacaact attgttaatg gcattgaagag  
 atctttctat  
 7501 gtctatgcaa atggaggccg ttggtcttgc aagactcaca attggaattg  
 tctcaattgt  
 7561 gacacatttt gcactggtag tacattcatt agtgatgaag ttgctcgtga  
 ttgtcacttc  
 7621 cagtttaaaa gaccaatcaa cctactgac cagtcactgt atattgttga  
 tagtgttgc  
 7681 gtgaaaaatg gcgcgttcca cctctacttt gacaaggctg gtaaaaagac  
 ctatgagaga  
 7741 catcgcctct cccattttgt caatttagac aatttgagag ctaacaacac  
 taaagggtca  
 7801 ctgcctatta atgtcatagt tttgatggc aagtcctaat gcgacgagtc  
 tgcctctaag  
 7861 tctgttcttg tgaactacag tcagctgatg tgccaacctt tctgttgc  
 tgaccaagct  
 7921 ctgtgatcag acgttgagga tagtactgaa gtttcgtta agatgttga  
 tgccttatgc  
 7981 gacacctttt cagcaacttt tagtgttctc atggaaaaac ttaaggcact  
 tgggtctaca  
 8041 gctcacagcg agttagcaaa ggggtgagct ttagatgggt tctttctac  
 attcgtgca  
 8101 gctgcccgac aagggtgtgt tgataccgat gttgacacaa aggatgttat  
 tgaatgtctc  
 8161 aaactttcac atcactctga cttagaagtg acaggtgaca gttgttaaca  
 tttcatgctc  
 8221 acctataata aggttgaaaa catgacgccc agagatcttg gcgcagtgt  
 tgactgtaat

8281 gcaaggcata tcaatgccca agtagcaaaa agtcacaatg ttctactcat  
 ctggaatgta  
 8341 aaagactaca tgtctttatc tgaacagctg cgtaaacaaa ttctagtgct  
 tgccaagaag  
 8401 aacaacatac ctcttagact aacttggtgt acaactagac aggttgctcaa  
 tgcataaact  
 8461 actaaaatct cactcaagggt tggtaagatt gttagtactt gttttaaact  
 tatgcttaag  
 8521 gccacattat tgtgcgttct tgcctgattg gtttgttata tegtattgac  
 agtacataca  
 8581 ttgtcaatcc atgatgggta cacaatatgaa atcattgggt acaaaagccat  
 tcaggatgggt  
 8641 gtccactcgtg acatcatctc tactgatgat tgttttgcaa ataaacatgc  
 tgggtttgac  
 8701 gcctgggttta gccacgtgtg tgggttcatac aaaaatgaca aaagctgcc  
 tctagtactc  
 8761 gctatcatta caagagagat tgggttcata gtgctgggtc tacgggttac  
 tctgctgaga  
 8821 gcaatcaatg gtgacttctt gcattttcta cctcgtgttt ttatgtgtgt  
 tggcaacatt  
 8881 tgcacacac ctcccaact cattgatgat agtgattttg ctacctctgc  
 ttgcgttctt  
 8941 gtgctgagt gtacaatttt taaggatgct atgggcaaac ctgtgccata  
 ttgttatgac  
 9001 actaatttgc tagaggggtc tatttcttat agtgagcttc gtccagacac  
 tctgtatgtg  
 9061 ctatggatg gttccatcat acagtttctt aacacttacc tggagggttc  
 tcttagagta  
 9121 gtaacaactt ttgatgctga gtactgtaga catggtacat gcgaagggtc  
 agaagtgggt  
 9181 atttgcctat ctaccagtgg tagatgggtt cttaataatg agcattacag  
 agctctatca  
 9241 ggagtttctt gtggtgttga tgcgatgaat ctcatagcta acatctttac  
 tctctttgtg  
 9301 caacctgtgg gtgctttaga tctgtctgct tcatgtatgg ctggtgttat  
 tattgccata  
 9361 ttgggtgactt gtgctgccta ctactttatg aaaltcagac gtgttttttg  
 tgagtaacaac  
 9421 catgtttgtg ctgctaattgc acttttgttt ttgatgtctt tcaactatact  
 ctgtctggta  
 9481 ccagcttaca gctttctgcc gggagctcac tcatgtcttt acttgacttt  
 gacattctat  
 9541 ttccacaatg atgtttcatt ctgtgctcac ctccaatggt ttgccatggt  
 ttctcttatt  
 9601 gtgccttttt ggataacacg aatctatgta ttctgtatct ctctgaagca  
 ctgcaattgg  
 9661 ttctttaaca actatcttag gaaaagagtc atgtttaatg gagttacatt  
 tagtaccttc  
 9721 gaggaggctg ctctgtgtac ctctttgctc aacaaggaaa tgcacttaaa  
 attgcgttag  
 9781 gagacactgt tgccacttac acagtataac aggtatcttg ctctatataa  
 caagtacaag  
 9841 tatctcagtg gacgcttaga tactaccagc tatcgtgaag cagcttgctg  
 ccacttagca  
 9901 aaggctctaa atgactttag caactcaggt gctgatgttc tctaccaaac  
 accacagaca  
 9961 tcaatcactt ctgctgtctc gcagagtggt tttaggaaaa tggcattccc  
 gtcaggcaaa  
 10021 gttgaagggt gcactgttaca agtaacctgt ggaactacaa ctcttaattg  
 attgtggttg  
 10081 gatgacacag tatactgtcc aagacatgct atttgacag cagaagacat  
 gcttaactct  
 10141 aactatgaag atctgctcat tgcgaatcc aacctagct ttctgtttca  
 ggtgggcaat  
 10201 gttcaacttc gtgttatttg ccattctatg caaaattgct tgcctagggt  
 taaagttagt  
 10261 acttctaacc ctaagacacc caagtataaa ttgttcgcta tccaacctgg  
 tcaaacattt  
 10321 tcagttctag catgctacaa tgggttcacca tctgggtgtt atcagtgctg  
 catgagacct  
 10381 aatcatacca ttaaagggtc ttctcttaat ggatcatgtg gtagtggttg  
 ttctaactat

10441 gattatgatt gctgtcttt ctgctatatg catcatatgg agcttccaac  
aggagtacac  
10501 gctggtagct acttagaagg taaattctat ggtccatttg ttgacagaca  
aactgcacag  
10561 gctgcaggta cagacacaac cataacatta aatgttttgg catggcgtga  
tgctgctgtt  
10621 atcaatgggt ataggtgggt tcttaataga ttcaccacta ctttgaatga  
ctttaacctt  
10681 gtggcaatga agtacaacta tgaacctttg acacaagatc atgttgacat  
attgggacct  
10741 ctttctgctc aaacaggaat tgcgctcta gatatgtgtg ctgctttgaa  
agagctgctg  
10801 cagaatggta tgaatggctg tactatcctt ggtagcacta ttttgaaga  
tgagtttaca  
10861 ccatttgatg ttgttagaca atgctctggg gttacccttc aaggtaagtt  
caagaaaatt  
10921 gttaaaggga ctcactattg gatgctttta actttcttga catcactatt  
gattcttgggt  
10981 caaagtacac agtggctact gttttctctt gtttacgaga atgctttctt  
gccatttact  
11041 cttgggtatta tggcaattgc tgcattgtct atgctgcttg ttaagcataa  
gcacgcatcc  
11101 ttgtgcttgt ttctgttacc ttctcttgca acagttgctt actttaatat  
ggtctacatg  
11161 cctgctagct ggggtgatgc tatcatgaca tggcttgaat tggtgacac  
tagctgtctt  
11221 gggtataggc ttaaggattg tgttatgtat gcttcagctt tagttttgct  
tattctcatg  
11281 acagctcgca ctgtttatga tgatgtctgt agacgtgttt ggacactgat  
gaatgtcatt  
11341 acacttggtt acaaagtcta ctatggtaat gctttagata aagctatttc  
catgtgggac  
11401 ttagttattt ctgtaacctc taactattct ggtgtcgtta cgactatcat  
gttttttagct  
11461 agagctatag tgtttgtgtg tgttagtat taccatttgt tatttattac  
tggcaacacc  
11521 ttacagtgtg tcatgcttgt ttattgttct ttaggcattt gttgtgctg  
ctactttggc  
11581 cttttctggt tactcaacgc ttacttcagg cttactcttg gtgtttatga  
ctacttggtc  
11641 tctacacaag aatttaggta tatgaactcc caggggcttt tgctctctaa  
gagtagtatt  
11701 gatgctttca agcttaacat taagttgttg ggtattggag gtaaacattg  
tatcaagggt  
11761 gctactgtac agtctaaaat gtctgacgta aagtgcacat ctgttggtact  
gctctcggtt  
11821 ctccaacaac tttagagtaga gtcatcttct aaattgtggg cacaatgtgt  
acaactccac  
11881 aatgatattc ttcttgcaaa agacacaact gaagcttttg agaagatggt  
ttctcttttg  
11941 tctgttttgc tatccatgca ggggtgctga gacattaata ggttgtgcga  
ggaaatgctc  
12001 gataaccgtg ctactcttca ggctattgct tcagaattta gttctttacc  
atcatatgcc  
12061 gcttatgcca ctgccaggga ggccatagag caggctgtag ctaatggtga  
ttctgaagtc  
12121 gttctcaaaa agttaagaa atctttgaat gtggctaaat ctgagtttga  
ccgtgatgct  
12181 gccatgcaac gcaagttgga aaagatggca gatcaggcta tgaccaaat  
gtacaaacag  
12241 gcaagatctg aggacaagag ggcaaaagta actagtcta tgcaacaat  
gctcttcaat  
12301 atgcttagga agcttgataa tgatgcactt aacaacatta tcaacaatgc  
gcgtgatggt  
12361 tgtgttccac tcaacatcat accattgact acagcagcca aactcatggt  
tgtgtccct  
12421 gattatggta cctacaagaa cacttgtgat ggtaacacct ttacatatgc  
atctgcactc  
12481 tgggaaatcc agcaagttgt tgatgcggat agcaagattg ttcaacttag  
tgaaattaac  
12541 atggacaatt caccaaaatt ggcttggcct cttatttgta cagctctaag  
agccaactca

12601 gctgttaaac tacagaataa tgaactgagt ccagtagcac tacgacagat  
 gtcctgtgcy  
 12661 gctggtacca cacaacacgc ttgtactgat gacaatgcac ttgcctacta  
 taacaatcgc  
 12721 aagggaggta ggtttgtgct ggcattacta tcagaccacc aagatctcaa  
 atgggctaga  
 12781 ttccctaaga gtgatggtac aggtacaatt tacacagaac tggaaaccacc  
 ttgtaggttt  
 12841 gttacagaca caccaaaagg gcttaagtg aaatacttgt acttcatcaa  
 aggtctaaac  
 12901 aacctaaata gaggtatggt gctgggcagt ttagctgcta cagtacgtct  
 tcaggctgga  
 12961 aatgctacag aagtacctgc caatcaact gtgcttctct tctgtgcttt  
 tgcagtagac  
 13021 cctgctaaag catataagga ttacctagca agtggaggac aaccaatcac  
 caactgtgtg  
 13081 aagatgttgt gtacacacac tggtagagga caggcaatta ctgtaacacc  
 agaagctaac  
 13141 atggaccaag agtcctttgg tggcgcttca tgttgtctgt attgtagatg  
 ccacattgac  
 13201 catccaatc ctaaaggatt ctgtgacttg aaaggttaagt acgtccaat  
 acctaccact  
 13261 tgtgctaagt acccagtggtg ttttacctt agaaacacag tctgtaccgt  
 ctgagggaatg  
 13321 tggaaagggt atggctgtag ttgtgaccaa ctccgcgaac ccttgatgca  
 gtctgaggat  
 13381 gcactaacgt ttttaacgg gtttgcggtg taagtgcagc cgtcttaca  
 cgtgaggga  
 13441 caggcactag tactgatgac gtctacaggc cttttgatat ttacaacgaa  
 aaagtgtgtg  
 13501 gttttgcaa gttcctaaaa actaattgct gtcgottcca ggagaaggat  
 gaggaaggga  
 13561 atttattaga cttctacttt gtatgtaaga ggcatactat gtctaactac  
 caacatgaag  
 13621 agactattta taacttggtt aaagattgtc cagcgggtgc tgcctatgac  
 ttttcaagt  
 13681 tttagtaga ttgtgacatg gtaccacata tatcaogtca cgtcttaact  
 aaatacacaa  
 13741 tggctgattt agtctatgct ctacgtcatt ttgatgaggg taattgtgat  
 acattaaaag  
 13801 aaatactcgt cacatacaat tgcgtgatg atgattattt caataagaag  
 gattggtatg  
 13861 acttcgtaga gaactcctgac atcttacgcg tatatgctaa cttagggtgag  
 cgtgtacgcc  
 13921 aatcattatt aaagactgta caattctgcg atgctatgcg tgatgcaggc  
 attgtaggcg  
 13981 tactgacatt agataatcag gatcttaagt ggaactggta cgaattcggg  
 gatttcgtac  
 14041 aagtagcacc aggcgtcgga gttcctattg tggattcata ttaactcattg  
 ctgatgccca  
 14101 tcttcacttt gactagggca ttggctgctg agtccatcat ggaatgctgat  
 ctgcgaatac  
 14161 caattattaa ttgggatttg ctgaaatag attttacgga agagagactt  
 tgcctcttcg  
 14221 accgttattt caaatattgg gaccagacat accatcccaa ttgtattaac  
 tgtttggatg  
 14281 atagggtgat ccttcattgt gcaaacctta atgtgttatt tctactgtg  
 tttccacctta  
 14341 caagtttttg accactagta agaaaaatat ttgtagatgg tgtctctttt  
 gtgtttcaa  
 14401 ctggatacca ttttcgtgag ttaggagtgc tacataatca ggaatgtaac  
 ttacatagct  
 14461 cgcgtctcag ttccaaggaa cttttagtgt atgctgtga tccagctatg  
 catgcagctt  
 14521 ctggcaattt atgtctagat aaacgcacta catgcttttc agtagctgca  
 ctaacaaaaca  
 14581 atgttgcttt tcaaacgtgc aaaccggta attttaataa agacttttat  
 gactttgctg  
 14641 tgtctaaaag tttctttaag gaagggaagt ctgttgaact aaaaacttc  
 tctttgtc  
 14701 aggatggcaa cgtgctatc agtgattatg actattatcg ttataatcg  
 ccaacaatgt

14761 gtgatatcag acaactccta tctgtagttg aagtgttga taataacttt  
 gattgttacg  
 14821 atggtggctg tattaatgcc aaccaagtaa tctgtaacaa tctggataaa  
 tcagctgggt  
 14881 tcccatltaa taatgggggt aaggctagac tttattatga ctcaatgagt  
 tatgaggatc  
 14941 aagatgcact ttctgcgtat actaagccta atgtcatccc tactataact  
 caaatgaatc  
 15001 ttaagtatgc cattagtcca aagaatagag ctgcaccgt agctgggtgc  
 tctatctgta  
 15061 gtactatgac aaatagacag ttctcatcaga aattattgaa gtaaatagcc  
 gccactagag  
 15121 gagctactgt ggttaattgga acaagcaagt tttacgggtg ctggcataat  
 atgttaaaaa  
 15181 ctgtttacag tgaatgtagaa actccacacc ttatgggttg ggattatcca  
 aaatgtgaca  
 15241 gagccatgcc taacatgctt aggataatgg cctctcttct tcttctctgc  
 aaacataaca  
 15301 cttgctgtaa cttatcacac cgtttctaca ggttagctaa cgaagtgtgcg  
 caagtattaa  
 15361 gtgagatggt catgtgtggc ggctcactat atgttaaacc aggtggaaca  
 tcactcgggt  
 15421 atgctacaac tgcttatgct aatagtgtct ttaacatttg tcaagctggt  
 acagccaatg  
 15481 taaatgcact tctttcaact gatggtaata agatagctga caagtatgtc  
 cgcactctac  
 15541 aacacaggct ctatgagtgt ctctatagaa atagggatgt tgatcatgaa  
 ttctgggatg  
 15601 agttttacgc ttacctgctt aaacatttct ccatgatgat tctttctgat  
 gatgcggtg  
 15661 tbtgctataa cagtaactat gcggtcgaag gtttagtagc tagcattaag  
 aactttaagg  
 15721 cagttcttta ttatcaaaat aatgtgttca tgcctgaggc aaaaatgttg  
 actgagactg  
 15781 acottactaa aggaacctac gaattttgct cacagcatac aatgctagtt  
 aaacaaggag  
 15841 atgattacgt gtacctgcct taccagatc catcaagaat attaggcgca  
 ggctgttttg  
 15901 tggatgatat tgtcaaaaca gatggtagac ttatgattga aagggtctgt  
 tcaactggcta  
 15961 ttgatgctta ccacttaca aaacatccta atcaggagta tgcgtgatgc  
 tttcaactgt  
 16021 atttacaata cattagaag ttacatgatg agcttactgg ccacatgttg  
 gacatgtatt  
 16081 ccgtaagtgt aactaatgat aacacctcac ggtactggga acctgagttt  
 tatggggcta  
 16141 tgtacacacc acatacagtc ttgcaggctg taggtgcttg tgtattgtgc  
 aattcacaga  
 16201 cttaactctg ttgcggtgct tgtattagga gaccattcct atgttgcaag  
 tgctgtctatg  
 16261 acctgtcat tcaacatca cacaactag tggttctgt taatccctat  
 gtttgcaatg  
 16321 cccaggttg tgaatgcact gatgtgacac aactgtatct agggaggtatg  
 agctattatt  
 16381 gcaagtcaca taagcctccc attagtttc cattatgtgc taatggctag  
 gtttttggtt  
 16441 tatacaaaaa cacatgtgta ggcagtgaca atgtcactga cttcaatgcy  
 atagcaaacat  
 16501 gtgattggac taatgctggc gattacatac ttgccaacac ttgtactgag  
 agactcaagc  
 16561 ttttctgcagc agaaacgctc aaagccactg aggaacatt taagctgtca  
 tatggatttg  
 16621 ccaactgtacg cgaagtactc tctgacagag aattgcactt tttatgggag  
 ttgggaaaac  
 16681 ctgagaccac attgaacaga aactatgtct ttaactggta ccgtgtaact  
 aaaaatagta  
 16741 aagtacagat tggagagtac acctttgaaa aaggtgacta tgggtgatgct  
 gttgtgtaca  
 16801 gaggtactac gacatacaag ttgaatgttg gtgattactt tgggttgaca  
 tctcacactg  
 16861 taatgccact tagtgcaact actctagtgc cacaagagca ctatgtgaga  
 attactggct

16921 tgtaccaaac actcaacatc tcagatgagt ttcttagcaa tggttgcaaat  
tatcaaaaagg  
16981 tcggcatgca aaagtactct acactccaag gaccacctgg tactggtaag  
agtcattttg  
17041 ccactggact tgctctctat taccatctg ctcgcatagt gtatacgga  
tgctctcatg  
17101 cagctgttga tgccctatgt gaaaaggcat taaaatattt gcccatagat  
aaatgtagta  
17161 gaatcatacc tgcgcgtgag cgcgtagagt gttttgataa attcaaatgt  
aatccaacac  
17221 tagaacagta tgttttctgc actgtaaatg cattgccaga aacaactgct  
gacattgtatg  
17281 tctttgatga aatctctatg gctactaatt atgacttgag tggttgcaat  
gctagacttc  
17341 gtgcataaaca ctacgtctat attggcgatc ctgctcaatt accagcccc  
cgacactgac  
17401 tgactaaaagg cacactagaa ccagaatatt ttaattcagt gtgcagactt  
atgaaaacaa  
17461 taggtccaga catgttcctt ggaacttgct gccgttgctc tgctgaaatt  
gttgacactg  
17521 tgaagtgttt agtttatgac aataagctaa aagcacacaa ggataagtca  
gctcaatgct  
17581 tcaaaatggt ctacaaagggt gttattacac atgatgttct atctgcaatc  
aacagacctc  
17641 aaataggcgt tgaagagaa ttctctacac gcaactctgc ttggagaaaa  
gctgttttta  
17701 tctacacctta taattcacag aacgtgtgag cttcaaaatc cttaggattg  
cctacgcaga  
17761 ctgttgatcc atcacagggt tctgaatatg actatgtcat attcacacaa  
actactgaaa  
17821 cagcacactc ttgtaatgct aacgccttca atgtggctat cacaaggcca  
aaaattggca  
17881 ttttgtcat aatgtctgat agagatcttt atgacaaact gcaatttaca  
agtctagaaa  
17941 taccacgtcg caatgtggct acattacaag cagaaaaatg aactggactt  
tttaaggact  
18001 gtagtaagat cattactggt ctctcatcta cacaggcacc tacacacctc  
agcggtgata  
18061 taaagttcaa gactgaagga ttatgtgttg acataccagg cataccaaaag  
gacatgacct  
18121 accgtagact catctctatg atgggtttca aaatgaatta ccaagtcaat  
ggttacccta  
18181 atatgtttat caccgcgcaa gaagctatcc gtcacgttcg tgcgtggatt  
ggctttgatg  
18241 tagagggtcg tcatgcaact agagatgctg tgggtactaa cctacctctc  
cagctaggat  
18301 ttctacacag ttgtaactta gtagctgtac cgactgggta tggttgacct  
gaaaataaca  
18361 cagaattcac cagagttaat gcaaaacctc caccaggtga ccagttttaa  
catcttatac  
18421 cactcatgta taaaggcttg ccttggaaatg tagtgctgat taagatagta  
caaatgctca  
18481 gtgatacact gaaaggattg tcagacagag tcgtgttcgt cctttggggc  
catggctttg  
18541 agcttacatc aatgaagtac ttgtcaaga ttggacctga aagaactgtg  
tgctctgttg  
18601 acaaacgtgc aactgtctt tctactcat cagatactta tgctctgttg  
aatcattctg  
18661 tgggttttga ctatgtctat aaccatttta tgattgatgt tcagcagtgg  
ggctttacgg  
18721 gtaaccttca gactaaocat gaccaacatt gccaggtaga tggaaatgca  
catgtggcta  
18781 gttgtgatgc tatcatgact agatgtttag cagtcctatg gtgctttgtt  
aagcgcgttg  
18841 attggtctgt tgaataacct attataggag atgaactgag ggtaatttct  
gcttcagaaa  
18901 aagtacaaca catggtttgt aagctctgcat tggctgtgta taagtttcca  
gttcttcatg  
18961 acattggaaa tccaaggctc atcaagtgtg tgcctcaggc tgaagttagaa  
tggaaagtct  
19021 acgatgctca gccatgtagt gacaaagctt acaaaataga ggaactcttc  
tattcttatg



19081 ctacacatca cgataaatc actgatgggtg ttgtgttgtt ttggaattgt  
 aacgttgatc  
 19141 gttaccaccg caatgcaatt gtgtgttaggt ttgacacaag agtcttgtca  
 aacttgaaact  
 19201 taccaggctg tgatggtggt agtttgtatg tgaataagca tgcattccac  
 actccagctt  
 19261 tcgataaaaag tgcatttact aatttaaagc aattgccttt cttttactat  
 tctgatagtc  
 19321 cttgtgagtc tcatggcaca caagtagtgt cggatattga ttatgttcca  
 ctcaaatctg  
 19381 ctactgtgat tacacgatgc aatttaggtg gtgctgttgg cagacaccat  
 gcaaataggt  
 19441 accgacagta ctgggatgca tataatatga tgattttgct tggatttagc  
 ctatgggattt  
 19501 acaacaattt tgatacttat aacctgtgga atacatttac caggttacag  
 agtttagaaa  
 19561 atgtggccta taatgttgtt aataaaggac acttttagtg acacgccgcg  
 gaagcacctg  
 19621 tttccatcat taataatgct gtttacacaa aggtagatgg tattgatgtg  
 gagatctttg  
 19681 aaaataagac aacacttctt gttatgtgtg catttgagct ttgggctaag  
 cgtaacatta  
 19741 aaacagtgcg agagattaag atactcaata atttgggtgt tgatattcgt  
 gctaatactg  
 19801 taatctggga ctacanaaga gaagcccgag cacatgtatc tacaataggt  
 gtctgcacaa  
 19861 tgactgacat tgccaagaaa cctactgaga gtgctgttgc ttcacttact  
 gtctgttggg  
 19921 atggtagagt ggaaggacag gttagacctt ttagaaacgc ccgtaatggt  
 gttttaataa  
 19981 cagaagggtc agtcaagggt ctaaacctt caaaggagcc agcacaagct  
 agcgtcaatg  
 20041 gagtcacatt aattggagaa tcagtaaaaa cacagtttaa ctactttaag  
 aaagtagacg  
 20101 gcattattca acngttgctt gaacotact ttactcagag cagagactta  
 gaggatttta  
 20161 agcccgatgc acaaatgaa actgaacttc tggagctgcg tatggatgaa  
 ttcatacagc  
 20221 gatataagct cgagggtcat gccttcgaac acatcggtta tggagatttc  
 agtcatggac  
 20281 aacttggcgg tcttcattta atgataggct tagccaagcg ctcaacaagt  
 tcaccactta  
 20341 aatttagagga ttttatccct atggacagca cagtgaaaaa ttacttcata  
 acagatgcgc  
 20401 aaacagggtc atcaaatgt gtgtgttctg tgatcgatct tctacttgat  
 gactttgtcg  
 20461 agataataaa gtcacaagat ttgtcagtga tttcaaaagt ggtcaagggt  
 acaattgact  
 20521 atgctgaaat ttcattcatg ctttgggtga aggatggaca tggttgaaac  
 tcttaccgaa  
 20581 aactacaagc aagtcaagcg tggcaaccag gtgtgcgat gcctaacttg  
 tacaagatgc  
 20641 aaagaatgct tcttgaanaa tgtgaccttc agaattatgg tgaanaatgct  
 gttataccaa  
 20701 aaggaataat gatgaatgct gcaaaagtata ctcaactgtg tcaacttata  
 aatacactta  
 20761 ctttagctgt accctacaac atgagagtta ttcactttgg tctgtggtct  
 gataaaggag  
 20821 ttgacaccag tacagctgtg ctacagcaat ggttgccaac tggcacacta  
 ctgtgcgatt  
 20881 cagatcttaa tgacttctgc tccgacgcag attctacttt aattggagac  
 tgtgcaacag  
 20941 tacatacggc taataaatgg gaccttatta ttagcgatat gtagtaccct  
 aggaccaaac  
 21001 atgtgacaaa agagaatgac tctaaagaag ggtttttcac ttactgtgtg  
 ggatttataa  
 21061 agcaaaaact agccctgggt ggttctatag ctgtaaatg atacagagcat  
 tcttgggaatg  
 21121 ctgaccttta caagcttatg ggccatttct catgggtggac agcttttgtt  
 acaaatgtaa  
 21181 atgcatcatc atcggaagca tttttaattg gggctaacta tcttggcaag  
 ccgaaggaaac

21241 aaattgatgg ctataccatg catgctaact acattttctg gaggaacaca  
 aatcctatcc  
 21301 agttgtcttc ctattcactc ttgacatga gcaaatctcc tcttaaatca  
 agagggaactg  
 21361 ctgtaatgtc tcttaaggag aatcaaatca atgatatgat ttattctctt  
 ctggaaaaag  
 21421 gtaggccttat cattagagaa aacaacagag ttgtggttc aagtatatt  
 ctgttlaaca  
 21481 actaaacgaa catgtttatt ttcttattat ttcttactct cactagtgtg  
 agtgaccttg  
 21541 accggtgcac cacttttgat gatgttcaag ctctcaatta cactcaacat  
 acttcatcta  
 21601 tgaggggggt ttactatcct gatgaaatt ttatgcaga cactctttat  
 ttaactcagg  
 21661 atttatttct tccattttat tctaagtta cagggtttca tactattaat  
 catacgtttg  
 21721 gcaaccctgt catacctttt aaggatggta ttattttgc tgccacagag  
 aaatcaaatg  
 21781 ttgtccgtgg ttgggttttt ggttctacca tgaacaacaa gtcacagtgc  
 gtgattatta  
 21841 ttaacaattc tactaatgtt gttatcagag catgttaact tgaattgtgt  
 gacaaccctt  
 21901 tcttgcctgt ttctaaacct atgggtacac agacacatc tatgatattc  
 gataatgcac  
 21961 ttaattgcac ttctcagatc atatctgatg ccttttcgct tgatgtttca  
 gaaaagtcag  
 22021 gtaattttta acacttaaga gagtttgtgt ttaaaaaata agatgggttt  
 ctctatgttt  
 22081 ataagggtca tcaacctata gatgtagttc gtgatctacc ttctggtttt  
 aacactttga  
 22141 aacctatttt taagtgtcct ctgggtatta acattacaaa ttttagagcc  
 attcttacag  
 22201 ccttttcacc tgctcaagac atttggggca cgtcagctgc agcctatttt  
 gttggctatt  
 22261 taaagccaac tacatttatg ctcaagtatg atgaaaatgg tacaatcaca  
 gatgctgttg  
 22321 attgttctca aaatccactt gctgaactca aatgctctgt taagagcttt  
 gagattgaca  
 22381 aagggaattta ccagacctct aatttcaggg ttgtccctc aggagatggt  
 gtgagattcc  
 22441 ctaattattc aaacttgtgt ccttttggag aggtttttta tgctactaaa  
 ttcccttctg  
 22501 tctatgcacg ggagagaaaa aaaatttcta attgtgttgc tgattactct  
 gtgctctaca  
 22561 actcaacatt tttttcaacc ttttaagtgt atggcggttc tgccactaag  
 ttgaatgac  
 22621 ttgtctctc caatgtctat gcagattctt ttgtagtcaa gggagatgat  
 gtaagacaaa  
 22681 tagcgcagg acaaaactgg gttattgctg attataatta taaattgcca  
 gatgatttca  
 22741 tgggttgtgt ccttgcctgg aatactagga acattgatgc tacttcaact  
 ggttaattca  
 22801 attataaata taggtatctt agacatggca agcttagggc ctttgagaga  
 gacatatcta  
 22861 atgtgccttt ctccctgat ggcaaacctt gaacccacc tgctctta  
 tgtattggc  
 22921 cattaaatga ttatggtttt tacaccacta ctggcatttg ctaccaacct  
 tacaagagttg  
 22981 tagtaacttc ttttgaactt ttaaatgcac cgcccaagggt ttgtggacca  
 aaattatcca  
 23041 ctgaccttat taagaaccag tgtgtcaatt ttaattttta tggactcact  
 ggtactgttg  
 23101 tgtaactcct ttcttcaaag agatttcaac catttcaaca atttggcgt  
 gatgtttctg  
 23161 atttactcga ttccgttcga gatcctaaaa catctgaaat attagacatt  
 tcaacttgcc  
 23221 cttttggggg tgtaagtgtt attacacctg gaacaaatgc ttactctgaa  
 gttgtgttcc  
 23281 tatatcaaga ttttaactgc actgatgttt ctacagcaat tcatgcagat  
 caactcaac  
 23341 cagcttgccg catatattct actggaaaca atgtattcca gactcaagca  
 ggtgtcttta

23401 taggagctga gcatgctgac acttcttatg agtgcgacat tctctatgga  
 gctggcattt  
 23461 gtgctagtta ccatacagtt tctttattac gtagtactag ccaaaaatct  
 attgtggctt  
 23521 atactatgtc tttaggtgct gatagttcaa ttgcttactc taataacacc  
 attgctatc  
 23581 cctaactt ttcaattagc attactacag aagtaatgcc tgtttctatg  
 gctaaaacct  
 23641 ccgtagattg taatatgtac atctgcggag attctactga atgtgcta  
 ttgcttctcc  
 23701 aatatggtag cttttgcaca caactaaac gtgcactctc aggtattgct  
 gctgaacagg  
 23761 atcgcaacac acgtgaagtg ttgctcaag tcaaacaaat gtacaaaacc  
 ccaactttga  
 23821 aatattttgg tggttttaat ttttcacaaa tattacctga cctctaaag  
 ccaactaaga  
 23881 ggtcttttat tgaggacttg ctcttaata aggtgacact cgctgatgct  
 ggtctcatga  
 23941 agcaatatgg cgaatgccta ggtgatatta atgctagaga tctcatttgt  
 ggcgaagaat  
 24001 tcaatggact tacagtgttg ccacctctgc tcaactgatg tatgattgct  
 gctactactg  
 24061 ctgctctagt tagtggtact gccactgctg gatggacatt tgggtgctgc  
 gctgctcttc  
 24121 aaataccttt tgcctatgcaa atggcatata ggttcaatgg cattggagtt  
 acccaaaatg  
 24181 ttctctatga gaacaaaaa caaatcgcca accaatttaa caaggcgatt  
 agtcaaatc  
 24241 aagaatcact tacaacaaca tcaactgcat tgggcaagct gcaagacgtt  
 gttaacccaga  
 24301 atgctcaagc attaaacaca cttgttaaac aacttagctc taattttggt  
 gcaatttcaa  
 24361 gtgtgctaaa tgatctctt tcgcgacttg ataaagtga ggcgcaggta  
 caaatgaca  
 24421 ggttaattac aggcagactt caaagccttc aaacctatgt aacacaacaa  
 ctaactcagg  
 24481 ctgctgaaat cagggtcttc gctaactctg ctgctactaa aatgtctgag  
 tgtgtctctg  
 24541 gacaatcaaa aagagttgac ttttggtaa agggctacca ccttatgtcc  
 ttccctcaag  
 24601 cagcccgca tgggtgtgct ttctacatg tcactgatgt gccatccag  
 gagaggact  
 24661 tcaccacagc gccagcaatt tgtcatgaag gcaagcata cttccctctg  
 gaagtgtt  
 24721 ttgtgtttaa tggcacttct tggtttatta cacagaggaa cttcttttct  
 ccacaataaa  
 24781 ttactacaga caatacattt gtctcaggaa atgtgatgt cgttatgtgc  
 atcataaca  
 24841 acacagttta tgatcctctg caacctgagc ttgactcatt caaagaagag  
 ctggacaagt  
 24901 acttcwaaaa tcatatcatc ccagatgttg atcttgccga catttcaggc  
 attaacgctt  
 24961 ctgtcgtcaa cattcaaaaa gaaattgacc gcctcaatga ggtcgctaaa  
 aattcaaatg  
 25021 aatcactcat tgacctcaa gaattgggaa aatatgagca atataaaaa  
 tggccttggt  
 25081 atgtttggct cggcttcatt gctggactaa ttgocatgt catggttaca  
 atctgctttt  
 25141 gtgctatgac tagttgttgc agttgctca aggggtgatg ctcttgtggt  
 tcttgcgtga  
 25201 agtttgatga ggaatgactc gagccagttc tcaagggtgt caaatatcat  
 tccacataaa  
 25261 cgaacttatg gatttgttta tgagattttt tactcttaga tcaattactg  
 cacagccagt  
 25321 aaaaattgac aatgctctc ctgcaagtac tgttcatgct acagcaacga  
 taccgtaca  
 25381 agcctcactc ctttcggat ggtctgttat tggcgttgca tttcttctg  
 tttttcagag  
 25441 cgctaccaaa ataattgcgc tcaataaaa atggcagcta gcccttata  
 agggcttcca  
 25501 gttcatttgc aatttactgc tgcatttgt taccatctat tcaactctt  
 tctgtctgc

25561 tgcaggatgtag gaggcgcaat ttttgacct ctatgecttg atatatatttc  
tacaatgcat  
25621 caacgcatgt agaattatta tgagatgttg gctttgttgg aagtgcgaat  
ccaagaaccc  
25681 attactttat gatgccaaact accttgtttg ctggcacaca cataactatg  
actactgtat  
25741 accatataac agtgtcacag atacaattgt cgttactgaa ggtgacggca  
tttcaacacc  
25801 aaaactcaaa gaagactacc aaattggttg ttattctgag gataggcact  
caggtgttaa  
25861 agactatgtc gttgtacatg gctatttcac cgaagtttac taccagcttg  
agtctacaca  
25921 aattactaca gacactggta ttgaaaatgc tacattcttc atctttaaca  
agcttgttaa  
25981 agaccacccg aatgtgcmaa tacacacaat cgacggctct tcaggagttg  
ctaactccagc  
26041 aatggatcca atttatgatg agccgacgac gactactagc gtgccttctg  
aagcacaaga  
26101 aagtgagtag gaacttatgt actcattcgt ttcggaagaa acaggtagct  
taaatgttaa  
26161 tagcgtactt ctttttcttg ctttcgtggt attcttgcta gtcacactag  
ccatccttac  
26221 tgcgcttcga ttgtgtgcgt actgctgcaa tattgttaac gtgagtttag  
taaaacccaac  
26281 gggttacgtc tactcgcgtg ttaaaaatct gaactcttct gaaggagttc  
ctgatcttct  
26341 gggtcaaacg aactaactat tattattatt ctgtttggaa ctttaacatt  
gcttatcatg  
26401 gcagacaacg gtactattac cgttgaggag ctttaacaaac tccttggaaac  
atgggaacct  
26461 gtaaataggtt tcctattctc agcctggatt atgttactac aatttgctta  
ttctaattcg  
26521 aacagggttt tgtacataat aaagcttgtt ttcctctggc tcttctggcc  
agtaaacactt  
26581 gcttgttttg tgctgtctgc tgtctacaga attaatggg tgactggcgg  
gattgagatt  
26641 gcaatggctt gtattgtagg ctgtgatggt cttagctact tcgttgcctc  
cttcaggctg  
26701 tttgctcgta ccgctcaat gtggtcattc aaoccagaaa caaacattct  
tctcaatgtg  
26761 cctctccggg ggacaattgt gaccagaccg ctcatggaaa gtgaacttgt  
cattggtgct  
26821 gtgatcattc gtggtcaact gcgaatggcc ggacactccc tagggcgctg  
tgacattaa  
26881 gacctgccaa aagagatcac tgtggctaca tcacgaacgc tttcttatta  
caaattagga  
26941 gcgtcgacgc gtgtaggcac tgattcaggt tttgctgcat acaacgccta  
ccgtattgga  
27001 aactataaat taaatacaga ccacgccgt agcaacgaca atattgcttt  
gctagtacag  
27061 taagtacaaa cagatgttcc atcttgttga cttccaggtt acaatagcag  
agatattgat  
27121 tatcattatg aggaatttca ggaattgctat ttggaattct gacgttataa  
taagtcaaat  
27181 agtgagacaa ttatttaagc ctctaactaa gaagaattat tcggagttag  
atgatgaaga  
27241 acctatggag tttagattac cataaaacga acatgaaaat tattctcttc  
ctgacattga  
27301 ttgtatttac atcttgccg ctatatcact atcaggagtg tgttagaggt  
acgaactgtac  
27361 tactaaaaga accttgccca tcagggaacat acgagggcaa ttcaccattt  
cacctcttg  
27421 ctgacataaa atttgcaact acttgcaact gcacacactt tgccttggct  
tgtgctgacg  
27481 gtactcgaca tacctatcag ctgctgcaa gatcagtttc accaaaactt  
ttcatcagac  
27541 aagaggaggt tcaacaagag ctctactcgc cactttttct cattgttgc  
gctctagat  
27601 ttttaatact ttgcttcacc attaagagaa agacagaatg aatgagctca  
ctttaattga  
27661 cttctatttg tgctttttag cttttctgct attccttgtt ttaataatgc  
ttattatatt

27721 ttggcttttca ctcgaaatcc aggatctaga agaaccctgt accaaagtct  
aaacgaacat  
27781 gaaacttctc attgttttga ctgtatttc tctatgcagt tgcattgca  
ctgtagtaca  
27841 gcgctgtgca tctaataaac ctcatgtgct tgaagatcct tgtaaggtag  
aacactaggg  
27901 gtaatactta tagcactgct tggctttgtg ctctaggaaa ggttttacct  
ttcatagat  
27961 ggcacactat ggttcaaaca tgcacaccta atgttactat caactgtcaa  
gatccagctg  
28021 ttgggtgcgt ttagctagg tgttggtacc ttcataaagg tcacaaact  
gctgcattta  
28081 gagacgtact tgtgtttta aataaacgaa caaattaaaa tgtctgataa  
tggaccccaa  
28141 tcaaaccaac gtatgtcccc ccgatttaca tttggtggac ccacagattc  
aactgacaat  
28201 aaccagaatg gaggacgcaa tggggcaagg ccaaacacgc gccgaccca  
aggttttacc  
28261 aataatactg cgtcttggtt cactgctctc actcagcatg gcaggaggaa  
acttagattc  
28321 cctcgaggcc agggcggtcc aatcaacacc aatagtgtgc cagatgacca  
aattggctac  
28381 taccgaagag ctaccgcagc agttcgtggt ggtgacgcca aaatgaaaga  
gctcagcccc  
28441 agatgglact tctattacct aggaactggc ccagaagcct cacttcccta  
cggcgctaac  
28501 aaagaaggca tegtatgggt tgcacttagg ggagccttga atacaccaa  
agaccacatt  
28561 ggcaccgcga atcctaataa caatgtgtcc accgtgtac aacttctcta  
aggaaacaaca  
28621 ttgccaaaag gcttctacgc agaggggaagc agaggcgcca gtcaagcttc  
ttctcgctcc  
28681 tcatcacgta gtgcgggtaa ttcaagaat tcaactcctg gcagcagtag  
gggaaattct  
28741 cctgctcgaa tggctagcgg aggtggtgaa actgccctcg cgctattgct  
gctagacaga  
28801 ttgaaccagc ttgagagcaa agtttctggt aaaggccaac aacaacaagg  
ccaaactgtc  
28861 actaagaaat ctgctgtgca ggcattctaa aagcctcgcc aaaaacgtac  
tgccacaaaa  
28921 cagtacaacg tcaactcaagc atttgggaga cgtgggtccag aacaaccca  
aggaaatttc  
28981 ggggaaccaag acctaatcag acaaggaaact gattacaac attggccgca  
aattgcacaa  
29041 tttgctccaa gtgcctctgc attctttgga atgtcacgca ttggcagtag  
agtcacacct  
29101 tcgggaacat ggtgacttta tcatggagcc attaaattgg atgacaaaga  
tccacaattc  
29161 aaagacaacg tcatactgct gaacaagcac attgacgcat acaaacatt  
cccaccaaca  
29221 gagcctaata aggacaaaaa gaaaagact gatgaagctc agccttggc  
gcagagacaa  
29281 aagaagcagc ccaatgtgac tcttcttctt cggcgtgaca tggatgattt  
ctccagacaa  
29341 ctccaaaatt ccatgagtgg agcttctgct gattcaactc aggcataaac  
actcatgatg  
29401 accacacaag gcagatgggc tatgtaaacg ttttcgcaat tccgtttacg  
atcatagtct  
29461 tactcttgtg cagaatgaat tctcgttaact aaacagacaa agtaggttta  
gttaacttta  
29521 atctcacata gcaatcttta atcaatgtgt aacattaggg aggacttgaa  
agagccacca  
29581 cattttcctc gagggcacgc ggaagtacat cgagggtaca gtgaataatg  
ctagggaagag  
29641 ctgcctatat ggaagagccc taatgtgtaa aattaatttt agtagtgcta  
tcccacgtg  
29701 attttaatat cttcttagga gaatgacaaa aaaaaaaaaa aaaaaaaaaa a

//

## SARS-associated Coronavirus

Search

Go!

About

News

Employment

Faculty

Contacts

Activities

Cancer Genetics

Advancing Platform

Functional Genomics

Gene Regulation

Gene Expression

Protein Modeling

Sequencing

Custom Services

Bioinformatics

Genome BC/Genome

Canada

Software

USBC

Training

Bioinformatics

Bioinformatics

Workshops

VanBUC

Resources

AcPrimer

CD/Gen

CD/Gen

ResearchBank.org

Genetic Code

Human Y-chromosome

Human Y-chromosome

Protein LAST

SARS Strain

Human BAC library

Additional Access

GEN: Internal web site

Attachment (2)

10/8/22, 3:03



More Information on SARS

- US Centre for Disease Control
- The World Health Organization

Download Sequence

The shotgun sequence reads of the TOR2 isolate are available for download.

### DOWNLOAD SEQUENCE

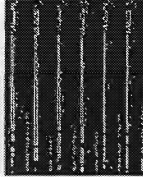
The fasta file contains 3,080 sequence reads totalling 3,621,777 base pairs.

The finished genome sequence assembly release 3, of the TOR2 isolate is available for download.

### DOWNLOAD SEQUENCE

The fasta file contains the first draft genome assembly totalling 29,751 base pairs. The assembly may contain errors.

Comparison to other Nidovirales



This image (2630a) was generated by the Sockeye viewer and shows the SARS TOR2 isolate's protein complement of the nidoviral family against build 7.4 of the SARS-

### Clone Availability

Clones derived from the SARS associated virus are now available from the SARS clone page.

### Genomic Sequence Availability

2003/04/29 The finished whole genome assembly is available for download. The fasta file contains the finished genome totalling 29,751 base pairs.

2003/04/14 The SARS-associated Coronavirus (TOR2) sequence has been deposited in the NCBI Genbank database under accession AY274119.

2003/04/12 The current genome shotgun data is available for download. The format is raw sequence reads in fasta format, and not vector or quality clipped.

2003/04/12 The draft whole genome assembly is available for download. The fasta file contains the first draft genome assembly totaling 29,756 base pairs. This assembly may contain errors.

### News Stories



The Linux Journal features a cover article about our lab's use of Linux in the sequencing of the SARS virus.

The cover of the magazine shows a 3-D image of a BLASTX comparison of the genome of the SARS TOR2 Coronavirus against the known proteins of other similar viruses. The image was generated using the Linux version of the Sockeye 3-D genome viewer, developed in our lab.

A feature story, "BC lab cracks suspected SARS code", is currently featured on cbc.ca

VANCOUVER - Scientists at a Canadian research centre say they've stopped the



genomic sequence of the virus believed to cause SARS, a move that could take the world a step closer to a reliable test for the illness. [more](#)

More information about SARS can be accessed [here](#).

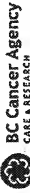
Text Size: A A A A

BC Cancer Agency news releases  
2003/04/12 | Genome Sciences Centre sequences SARS associated corona virus. [more](#)  
2003/04/07 | Genome Sciences Centre in partnership to sequence the SARS virus genome. [more](#)

Participants



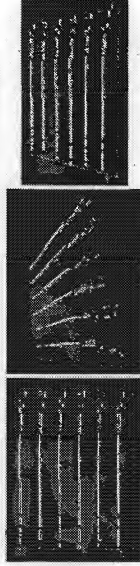
Logos of participating organizations.



BC Cancer Agency news releases  
2003/04/12 | Genome Sciences Centre sequences SARS associated corona virus. [more](#)  
2003/04/07 | Genome Sciences Centre in partnership to sequence the SARS virus genome. [more](#)

## Comparison to other Nidovirales

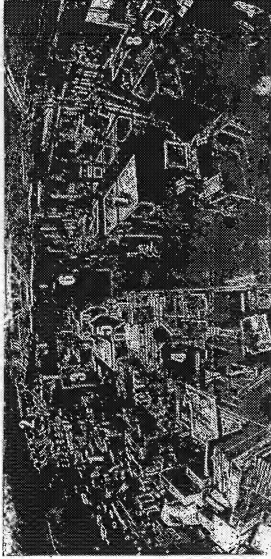
These images were generated by the Sackey genome viewer, developed at the Genome Sciences Centre. The images show the comparison of the sequence of the TOR2 isolate to the sequence of other related nidovirales. The tracks show blastx comparison to (1) avian infectious bronchitis virus, (2) bovine coronavirus, (3) human coronavirus 229E, (4) murine hepatitis virus, (5) porcine epidemic diarrhoea virus, (6) transmissible gastroenteritis virus. An annotated image is also available.



Images courtesy of the GSC bioinformatics team.

## Around the Lab

The Genome Sciences Centre carries a wide range of basic and applied research in genomics, including sequencing, physical mapping, functional genomics, cancer genomics, gene expression and technology development. The photo below shows the sequencing area of the lab. Other images are available [here](#).



1 96- and 384-well liquid handling robotics automate preparation of sequencing reactions.

2 The lab's power and modularity requirements are unique. Network and power connections drop down from the ceiling to facilitate arrangement of instruments.

3 DNA and other samples are stored in large freezers which maintain a constant temperature of -80 degrees Celsius.

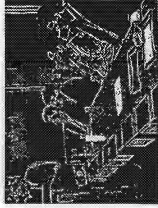
4 Thermocyclers are used to activate sequencing reactions by exposing the samples to a specific cycle of heating and cooling.

5 Robotic colony pickers are used to isolate and collect bacterial colonies which are used to amplify the DNA to be sequenced.

6 Linear-based Bioeruf cluster used for data analysis and sequence assembly builds.

7 State-of-the-art robotic sequencers include the ABI 3700, ABI 3730XL, Megabace 1000 and Megabace 4000. Sequencers are loaded with multiple 96- or 384-well plates and are capable of producing over 300,000 quality base pairs from a 384-well plate per run.

8 96- and 384-head liquid handling robots used to dispense reaction liquids to multiple plates.



2003/04/14 | Images from today's media conference





Images of people and equipment around the lab.

Contacts Email: [webmaster@bcgsc.ca](mailto:webmaster@bcgsc.ca) Phone: 604 877-6084

Powered by Plans

gb AAV37392.1	spike glycoprotein [SARS coronavirus B033]	2563	0.0
gb AAV3466.1	spike glycoprotein [SARS coronavirus civet019]	2581	0.0
gb AAV37395.1	spike glycoprotein [SARS coronavirus B012]	2560	0.0
gb AAV32009.1	spike glycoprotein [SARS coronavirus C025]	2569	0.0
gb AAV31984.1	spike glycoprotein [SARS coronavirus A001]	2560	0.0
gb AAV34632.1	spike glycoprotein [SARS coronavirus civet010]...	2562	0.0
gb AAV34631.1	spike glycoprotein [SARS coronavirus C028]	2555	0.0
gb AAV37328.1	spike glycoprotein [SARS coronavirus C018]	2563	0.0
gb AAV34646.1	spike glycoprotein [SARS coronavirus civet007]	2557	0.0
gb AAV31821.1	spike glycoprotein [SARS coronavirus A022] >gb...	2556	0.0
gb AAV3703.1	spike glycoprotein [recombinant coronavirus]	2223	0.0
gb F1P_00136139.1	spike protein [Bat SARS CoV Rm1/2004] >gb...	2062	0.0
gb Q95472.1	RecName: Full-Spike glycoprotein; Sho...	2961	0.0
gb F1P_00136134.1	spike protein [Bat SARS CoV Rp3/2004] >sp...	2056	0.0
gb F1P_00136235.1	spike glycoprotein [Bat coronavirus HKU3]...	2043	0.0
gb AAV31322.1	spike glycoprotein [bat SARS coronavirus HKU3-2]	2047	0.0
gb AAV31300.1	spike protein [Bat coronavirus (BtCoV/273/2005)]	2009	0.0
gb F1P_00136135.1	spike protein [Bat SARS CoV Rf1/2004] >gb...	2006	0.0
gb AAV32326.1	spike glycoprotein [SARS coronavirus GZ0403]	1554	0.0
gb AAV37226.1	S1 protein [SARS coronavirus GD322]	1460	0.0
gb AAV34630.1	S2 protein [SARS coronavirus GD322]	1103	0.0
gb AAV34644.1	truncated SARS coronavirus spike glycoprotein ...	1013	0.0
gb AAV34761.1	spike glycoprotein [SARS coronavirus TW-HP1] >...	927	0.0
gb AAV34787.1	spike glycoprotein [SARS coronavirus TW-YM2]	926	0.0
gb AAV34772.1	spike glycoprotein [SARS coronavirus TW-GD3] >...	921	0.0
gb AAV34781.1	spike glycoprotein [SARS coronavirus TW-GD4]	920	0.0
gb AAV31745.1	spike glycoprotein [SARS coronavirus CUHK-L2]	755	0.0
gb AAV30919.1	Spike glycoprotein [Bat coronavirus HKU9-2]	627	0.0
gb F1P_00136130.1	spike glycoprotein [Bat coronavirus HKU9]...	632	1e-178
gb AAV30927.1	Spike glycoprotein [Bat coronavirus HKU9-3]	622	6e-176
gb AAV30925.1	Spike glycoprotein [Bat coronavirus HKU9-4]	606	6e-171
gb F1P_00136131.1	spike glycoprotein [Bat coronavirus HKU5]...	591	2e-166
gb F1P_72226.1	spike protein [Bat coronavirus (BtCoV/133/20...]	583	3e-164
gb AAV34657.1	spike glycoprotein [SARS coronavirus cw049]	580	3e-163
gb AAV34654.1	Spike glycoprotein [Bat coronavirus HKU5-2]	579	7e-163
gb AAV30933.1	Spike glycoprotein [Bat coronavirus HKU5-3] >g...	578	2e-162
gb AAV31325.1	spike [Bat coronavirus (BtCoV/A434/2005)]	572	2e-161
gb AAV30937.1	Spike glycoprotein [Bat coronavirus HKU4-3]	572	9e-161
gb AAV30949.1	Spike glycoprotein [Bat coronavirus HKU4-2]	572	1e-160
gb F1P_00136133.1	spike glycoprotein [Bat coronavirus HKU4]...	572	1e-160

Alignments Select All Get selected sequences Distance tree of results Multiple alignment

>ref|NP\_028851.1| **G** E2 glycoprotein precursor [SARS coronavirus]  
 gb|AAP41037.1| spike glycoprotein [SARS coronavirus Tor2]  
 Length=1255

GENE ID: 1489668 S | E2 glycoprotein precursor [SARS coronavirus]  
 (Over 10 PubMed links)

Score = 2602 bits (6745), Expect = 0.0, Method: Compositional matrix adjust.  
 Identities = 1255/1255 (100%), Positives = 1255/1255 (100%), Gaps = 0/1255 (0%)

Query	1	MFIFLFLTLTSGSGLDRCTTFDDVQAPNYTQHTSSMRGVVYPDEIFRSDTLVLTQDLFL	60
Sbjct	1	MFIFLFLTLTSGSGLDRCTTFDDVQAPNYTQHTSSMRGVVYPDEIFRSDTLVLTQDLFL	60
Query	61	PFYSNVTFGHTINHTFGNPVIFPKDGIYFAATEKSNVVRGWVPGSTMMNKSQSVIIINNS	120
Sbjct	61	PFYSNVTFGHTINHTFGNPVIFPKDGIYFAATEKSNVVRGWVPGSTMMNKSQSVIIINNS	120
Query	121	TNNVIRACNFELCDNPPFAVSKPMGTQHTMTIFDPAFNCTFEYISDAPSLDVSEKSGNFK	180
Sbjct	121	TNNVIRACNFELCDNPPFAVSKPMGTQHTMTIFDPAFNCTFEYISDAPSLDVSEKSGNFK	180
Query	181	HLREFVFKNGDGLFYVYKGYPIDVVRDLPSPGNTLKPILFKLPLGINITNFRAILTAFSP	240

10/822,303  
 Attachment (3)

SEQIDNO:6042 vs

Tor2

Sbjct	181	HLREFVFNKNDGFLYVYKGYQPIDVVRDLPSGFNTLKP1FKLPLGINITNFRAILTAFSP HLREFVFNKNDGFLYVYKGYQPIDVVRDLPSGFNTLKP1FKLPLGINITNFRAILTAFSP	240
Query	241	AQDWTGTSAAAYFVGYLKPTTFMLKYDENGTTTDAVDCSQNPFLAEKCKSVKSPFEIDKGIY AQDWTGTSAAAYFVGYLKPTTFMLKYDENGTTTDAVDCSQNPFLAEKCKSVKSPFEIDKGIY	300
Sbjct	241	AQDWTGTSAAAYFVGYLKPTTFMLKYDENGTTTDAVDCSQNPFLAEKCKSVKSPFEIDKGIY AQDWTGTSAAAYFVGYLKPTTFMLKYDENGTTTDAVDCSQNPFLAEKCKSVKSPFEIDKGIY	300
Query	301	QTSNPRVPSGDVVRFPNITNLCPPGGEVFNATKFPSSVYAWERKKISNCVADYSVLYNSTF QTSNPRVPSGDVVRFPNITNLCPPGGEVFNATKFPSSVYAWERKKISNCVADYSVLYNSTF	360
Sbjct	301	QTSNPRVPSGDVVRFPNITNLCPPGGEVFNATKFPSSVYAWERKKISNCVADYSVLYNSTF QTSNPRVPSGDVVRFPNITNLCPPGGEVFNATKFPSSVYAWERKKISNCVADYSVLYNSTF	360
Query	361	FSTPKCYGVSATKLNLCFNSVYADSFVVGDDVRQIAPGQGTGVADYNYKLPDDFMGCV FSTPKCYGVSATKLNLCFNSVYADSFVVGDDVRQIAPGQGTGVADYNYKLPDDFMGCV	420
Sbjct	361	FSTPKCYGVSATKLNLCFNSVYADSFVVGDDVRQIAPGQGTGVADYNYKLPDDFMGCV FSTPKCYGVSATKLNLCFNSVYADSFVVGDDVRQIAPGQGTGVADYNYKLPDDFMGCV	420
Query	421	LAWNTRNIDATSTGNYNYKYRYLRHGKLRPPERDISNVFPSPDGKPCPTPALNCYWLND LAWNTRNIDATSTGNYNYKYRYLRHGKLRPPERDISNVFPSPDGKPCPTPALNCYWLND	480
Sbjct	421	LAWNTRNIDATSTGNYNYKYRYLRHGKLRPPERDISNVFPSPDGKPCPTPALNCYWLND LAWNTRNIDATSTGNYNYKYRYLRHGKLRPPERDISNVFPSPDGKPCPTPALNCYWLND	480
Query	481	YGYFTTTGIGYQPYRVVLSFELNAPATVCGPKLSTDLIKNCQVNFNFNGLTGTGVLT YGYFTTTGIGYQPYRVVLSFELNAPATVCGPKLSTDLIKNCQVNFNFNGLTGTGVLT	540
Sbjct	481	YGYFTTTGIGYQPYRVVLSFELNAPATVCGPKLSTDLIKNCQVNFNFNGLTGTGVLT YGYFTTTGIGYQPYRVVLSFELNAPATVCGPKLSTDLIKNCQVNFNFNGLTGTGVLT	540
Query	541	SSKRFQPPQFGPRDVSDFDTSVRDPKTSIILDISPCAFGGVSITPGTNASSEVAVLYQD SSKRFQPPQFGPRDVSDFDTSVRDPKTSIILDISPCAFGGVSITPGTNASSEVAVLYQD	600
Sbjct	541	SSKRFQPPQFGPRDVSDFDTSVRDPKTSIILDISPCAFGGVSITPGTNASSEVAVLYQD SSKRFQPPQFGPRDVSDFDTSVRDPKTSIILDISPCAFGGVSITPGTNASSEVAVLYQD	600
Query	601	VNCTDVSSTAIAHADQTPAWRIYSTGNVFPQTQAGCLIGAEHVDTSYECDPIGAGICASY VNCTDVSSTAIAHADQTPAWRIYSTGNVFPQTQAGCLIGAEHVDTSYECDPIGAGICASY	660
Sbjct	601	VNCTDVSSTAIAHADQTPAWRIYSTGNVFPQTQAGCLIGAEHVDTSYECDPIGAGICASY VNCTDVSSTAIAHADQTPAWRIYSTGNVFPQTQAGCLIGAEHVDTSYECDPIGAGICASY	660
Query	661	HTVSLRLSTSQKSIVAYTMSLGADSSIAYSNNTIAIPNFSISITTEVMPVMSAKTSVDC HTVSLRLSTSQKSIVAYTMSLGADSSIAYSNNTIAIPNFSISITTEVMPVMSAKTSVDC	720
Sbjct	661	HTVSLRLSTSQKSIVAYTMSLGADSSIAYSNNTIAIPNFSISITTEVMPVMSAKTSVDC HTVSLRLSTSQKSIVAYTMSLGADSSIAYSNNTIAIPNFSISITTEVMPVMSAKTSVDC	720
Query	721	NMYICGDSSTECANLLQYGSFCTQLNRLASGIAAEQDRNTRBVFPAQVKMYKPTPLKYFG NMYICGDSSTECANLLQYGSFCTQLNRLASGIAAEQDRNTRBVFPAQVKMYKPTPLKYFG	780
Sbjct	721	NMYICGDSSTECANLLQYGSFCTQLNRLASGIAAEQDRNTRBVFPAQVKMYKPTPLKYFG NMYICGDSSTECANLLQYGSFCTQLNRLASGIAAEQDRNTRBVFPAQVKMYKPTPLKYFG	780
Query	781	GFNFSQILPDLPLKTKRSFIEDLLFNKVTADAGFMKQYGECLGDINARDLICAKQFNGL GFNFSQILPDLPLKTKRSFIEDLLFNKVTADAGFMKQYGECLGDINARDLICAKQFNGL	840
Sbjct	781	GFNFSQILPDLPLKTKRSFIEDLLFNKVTADAGFMKQYGECLGDINARDLICAKQFNGL GFNFSQILPDLPLKTKRSFIEDLLFNKVTADAGFMKQYGECLGDINARDLICAKQFNGL	840
Query	841	TVLPPLLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQAYRPNFGIGVTQNVLYE TVLPPLLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQAYRPNFGIGVTQNVLYE	900
Sbjct	841	TVLPPLLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQAYRPNFGIGVTQNVLYE TVLPPLLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQAYRPNFGIGVTQNVLYE	900
Query	901	NQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFATSSVNL NQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFATSSVNL	960
Sbjct	901	NQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFATSSVNL NQKQIANQFNKAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFATSSVNL	960
Query	961	DLSRLDKVEAEVQIDRLITGRQLSQTQVTVQQLIRAAEIRASANLAATKMSCEVLGQSK DLSRLDKVEAEVQIDRLITGRQLSQTQVTVQQLIRAAEIRASANLAATKMSCEVLGQSK	1020
Sbjct	961	DLSRLDKVEAEVQIDRLITGRQLSQTQVTVQQLIRAAEIRASANLAATKMSCEVLGQSK DLSRLDKVEAEVQIDRLITGRQLSQTQVTVQQLIRAAEIRASANLAATKMSCEVLGQSK	1020
Query	1021	RVDPCGKGHYLMSFPQAAHPGVFLHVTVVPSQERNFTTAPACHGKAYFPREGVVFVN RVDPCGKGHYLMSFPQAAHPGVFLHVTVVPSQERNFTTAPACHGKAYFPREGVVFVN	1080
Sbjct	1021	RVDPCGKGHYLMSFPQAAHPGVFLHVTVVPSQERNFTTAPACHGKAYFPREGVVFVN RVDPCGKGHYLMSFPQAAHPGVFLHVTVVPSQERNFTTAPACHGKAYFPREGVVFVN	1080
Query	1081	GTSWFITQRNFPSPQIIITTDNTFVSGNCDVIGIINNTVVDPLQPELDSFKEBLDKYFKN GTSWFITQRNFPSPQIIITTDNTFVSGNCDVIGIINNTVVDPLQPELDSFKEBLDKYFKN	1140
Sbjct	1081	GTSWFITQRNFPSPQIIITTDNTFVSGNCDVIGIINNTVVDPLQPELDSFKEBLDKYFKN GTSWFITQRNFPSPQIIITTDNTFVSGNCDVIGIINNTVVDPLQPELDSFKEBLDKYFKN	1140
Query	1141	HTSPDVLGDLSIGINASVUNIQKSIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWL HTSPDVLGDLSIGINASVUNIQKSIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWL	1200
Sbjct	1141	HTSPDVLGDLSIGINASVUNIQKSIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWL HTSPDVLGDLSIGINASVUNIQKSIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWL	1200
Query	1201	GFIAGLIAVMVITILLCCMTSCCSCLKGACSCGSCCKDEDDSEPLVKGVKLHYT 1255 GFIAGLIAVMVITILLCCMTSCCSCLKGACSCGSCCKDEDDSEPLVKGVKLHYT 1255	
Sbjct	1201	GFIAGLIAVMVITILLCCMTSCCSCLKGACSCGSCCKDEDDSEPLVKGVKLHYT 1255 GFIAGLIAVMVITILLCCMTSCCSCLKGACSCGSCCKDEDDSEPLVKGVKLHYT 1255	

>gb|AAR07630.1| spike glycoprotein [SARS coronavirus BJ302]  
Length=1255

Score = 2600 bits (6739), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 1253/1255 (99%), Positives = 1254/1255 (99%), Gaps = 0/1255 (0%)

Query	1	MFIFLLFLTLTSGSGLDRCTTFDDVQAPNYTQHTSSMRGVVYDDEIFRSDTLVLTQDLFL	60
Sbjct	1	MFIFLLFLTLTSGSGLDRCTTFDDVQAPNYTQHTSSMRGVVYDDEIFRSDTLVLTQDLFL	60
Query	61	PFYSNVTGFHTINHTFGNEVPIPKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIILNNS	120



[CDC Home](#)

[Search](#)

[Health Topics A-Z](#)

**Office of Enterprise Communication**

**Phone (404)639-3286**

**MEDIA  
RELATIONS**

[Media Home](#) | [Contact Us](#)

## Press Release

*Attachment (4)  
10/822,303*

For Immediate Release

### Media Relations Links

- [About Us](#)
- [Contact Us](#)
- [Frequently Asked Questions](#)
- [Media Site Map](#)

### CDC News

- [Press Release Library](#)
- [Transcripts](#)
- [MMWR Summaries](#)
- [B-Roll Footage](#)
- [Upcoming Events](#)

### Related Links

- [Centers at CDC](#)
- [Data and Statistics](#)
- [Health Topics A-Z](#)
- [Image Library](#)
- [Publications, Software and Other Products](#)
- [Global Health Odyssey](#)

Find your [state](#) or [local](#) health department

[HHS News](#)

[National Health Observances](#)



Div. of Media Relations  
1600 Clifton Road  
MS D-14  
Atlanta, GA 30333  
(404) 639-3286  
Fax (404) 639-7394

April 14, 2003

Contact: CDC Media Relations:  
**404-639-3286**

## CDC Lab Sequences Genome of New Coronavirus

The Centers for Disease Control and Prevention (CDC) announced today that it has sequenced the genome for the coronavirus believed to be responsible for the global epidemic of severe acute respiratory syndrome or SARS. The CDC sequence is nearly identical to that determined by a Canadian laboratory late last week. The significant difference is that the CDC-determined sequence has 15 additional nucleotides, which provides the important beginning of the sequence, CDC scientists said.

The results came just 12 days after a team of 10 scientists, supported by numerous technicians, began working around the clock to grow cells taken from a throat culture taken from one of the SARS patients in Vero cells (African green monkey kidney cells) in order to reproduce the ribonucleic acid (RNA) of the disease-causing coronavirus. The new sequence has 29,727 nucleotides, which places it well within the typical RNA boundaries for coronaviruses. Members of this viral family tend to have between 29,000 and 31,000 nucleotides.

Identifying the genetic sequence of a new virus is important to efforts to treat or prevent it, said Dr. Julie Gerberding, CDC director. "Research laboratories can use this information to begin to target antiviral drugs, to form the basis for developing vaccines, and to develop diagnostic tests that can lead to early detection."

In sequencing the genome, CDC scientists worked closely with coronavirus experts at academic institutions across the United States. "This is an active, working community of scientific experts who have been contributing their knowledge and expertise throughout this investigation," said William Bellini, Ph.D., SARS laboratory team coordinator.

The nearly identical findings in the US and Canada are important because they were derived from different individuals who were infected in different countries. This suggests that the virus probably originated from a common source.

The CDC's analysis of the virus is far from finished, officials emphasized. Because coronaviruses have the ability to mutate rapidly, scientists will compare the sequences from viruses isolated in cell culture to those obtained from diseased tissues taken from SARS patients. "This is essentially a draft. Now we need to see if what we have identified in the laboratory matches what's causing disease in patients," Bellini said.

But the groundbreaking work of isolating the genomic sequence speeds

the task of comparison.

###

*CDC protects people's health and safety by preventing and controlling diseases and injuries; enhances health decisions by providing credible information on critical health issues; and promotes healthy living through strong partnerships with local, national, and international organizations.*

---

[Media Home Page](#) | [Accessibility](#) | [Privacy Policy](#) | [Contact Us](#)

[CDC Home](#) | [Search](#) | [Health Topics A-Z](#)

This page last updated April 14, 2003

URL: <http://www.cdc.gov/od/oc/media/pressrel/r030414.htm>

United States Department of Health and Human Services  
Centers for Disease Control and Prevention  
Office of Communication  
Division of Media Relations